DOCTOR OF PHILOSOPHY

Guitar Effects-Pedal Emulation and Identification

Holmes, Ben

Award date:
2019

Awarding institution:
Queen's University Belfast

Guitar Effects-Pedal Emulation and Identification

Benjamin Charles Holmes

BSc

Sonic Arts Research Centre

School of Electronics, Electrical Engineering and Computer Science

Queen's University Belfast

Submitted for the Degree of Doctor of Philosophy

May 2019

Acknowledgements

I would like to first thank my supervisor, Dr Maarten van Walstijn, for his endless support in not only completing this thesis but all of the hurdles along the way. I am also thankful for the supervision of Professor Kang Li and Dr Stuart Ferguson, their guidance has been a great help in shaping my PhD.

The Sonic Arts Research Centre at Queen's University Belfast has proved to be an exciting and welcoming community to join, and I am thankful for my time spent there. In particular I would like to thank my colleagues, Sandor Mehes and Jamie Bridges, with whom I shared a corner of the office and both their unique senses of humour and tastes in music.

I extend sincere gratitude to my examiners Professor Roger Woods and Dr Andrew McPherson for their thorough treatment of my thesis, engaging viva voce debate, and feedback to improve my work to its current state.

I am very grateful to Dr Martin Holters, whose shared passion for VA resulted in him hosting me at Helmut Schmidt University for a research trip to get to the bottom of germanium BJTs. I also owe thanks to the SEEECS technical support staff, namely Mr Jim Norney and Mr Gerry Rafferty, whom repeatedly helped me even after multiple cases of annihilated equipment.

I would not have attempted a PhD had it not been for the encouragement from my dad, and could not have completed it without the support of my mum. Je voudrais remercier ma femme formidable qui a toujours été a l'ecoute, ainsi que très patiente avec moi. For all the support of my family and friends I am deeply thankful.

The research presented in this thesis was made possible by the funding of the Department for Employment and Learning, Northern Ireland.

Abstract

A guitarist's search for 'tone' – their ideal timbre – can lead them to exploring countless guitar effects. Many players yearn for the tone of equipment from the early days of electric guitars. Now quickly becoming a mature field, Virtual Analogue modelling aims to digitally emulate analogue audio equipment in real-time, making potentially rare devices more accessible. A subset of Virtual Analogue of particular interest is found in the physical modelling of audio circuits, where circuit-level models are built from models of electronic components, drawing upon both widely transferable physical concepts and engineering methods.

Despite the increasing ubiquity of physical models, the simulated input/output behaviour is rarely compared to that of real circuits. The main contribution of this work is to reconcile this disparity with the presentation of two complementary identification procedures that aim to find a model with minimal difference to a reference circuit. Focusing on guitar pedals, measurements of the circuit are taken solely from existing input and output connections to reduce the required measurement time in comparison to measuring each component individually, which also prevents any damage being caused from the deconstruction of the device.

The identification approaches proposed in this study utilise an optimisation algorithm that minimises the difference between the output of candidate models to that of a reference circuit by modifying the values of the physical component parameters. Within the required simulation, the solving of nonlinear equations is a likely source of inefficiency and even failure, prompting the search for an algorithm that avoids these issues. Uncertainty about the accuracy of less well understood components can also lead to difficulties in the circuit identification. A component that is found to be markedly different is the germanium BJT – a core component present in a vintage case study – and is thus the focus of a component-level identification.

Of the two proposed identification procedures, the first aims only to minimise the output error, discarding accuracy at a component level, and placing a focus on minimising the computational expense of the identification. In addition to high fidelity models,

results point towards a strategy to overcome the curse of dimensionality when addressing circuits with a large number of components. The second, more physically valid procedure aims to retrieve accurate parameter values of each of the circuit's components such that the estimated component values remain valid under modifications to the circuit. To address possible non-convergence problems, an approach is developed that makes use of multiple measurement sets involving additional components of known value, thus introducing further constraints on the search space. The performance of both procedures is exemplified and evaluated by means of case studies.

Contents

List of Figures

List of Tables

List of Symbols

β_{f}	Forward current gain
β_{f}	Reverse current gain
$\boldsymbol{\kappa}$	Exhaustive summary containing each of the combinations of parameters in a model
$\boldsymbol{\theta}$	Parameter set consisting of physical component parameters of a circuit model
ϵ	Error between data sets
$\mathbf{A},\ \mathbf{B},\ \mathbf{C}$	State space matrices for the state update equation: past state, voltage inputs, and nonlinearities
$\mathbf{D}_{\mathrm{N}}.\ \mathbf{E}_{\mathrm{N}},\ \mathbf{F}_{\mathrm{N}}$	State space matrices for the nonlinear equation: past state, voltage inputs, and nonlinearities
$\mathbf{D}_{\mathrm{O}},\ \mathbf{E}_{\mathrm{O}},\ \mathbf{F}_{\mathrm{O}}$	State space matrices for the output equation
$\mathbf{G}_{\mathrm{R}},\ \mathbf{G}_{\mathrm{C}}$	Diagonal matrices containing conductances of resistors and capacitors
$\mathbf{i}$	Vector of current sources used in a Modified Nodal Analysis model
$\mathbf{i}_{\mathrm{v}}$	Vector of currents through voltage sources.
$\mathbf{J}$	Jacobian matrix containing first order partial derivatives of a vector valued function
$\mathbf{N}_{\mathrm{R}},\ \mathbf{N}_{\mathrm{C}}, \mathbf{N}_{\mathrm{N}}$	Incidence matrices of resistors, capacitors, and nonlinear components
$\mathbf{N}_{\mathrm{U}},\ \mathbf{N}_{\mathrm{O}}$	Incidence matrices of input voltage sources and outputs

$\mathbf{S}$	System matrix used in Modified Nodal Analysis
$\mathbf{u}$	State space model input, vector of voltage sources.
$\mathbf{x}$	State space model state
$\mathbf{y}$	State space model output
μ, μ^*	Mean, and mean of absolute values
ω	Angular frequency in radians, $\omega = 2\pi f$
ϕ	Signal phase
σ	Standard deviation
$\xi(\theta)$	Objective function for given parameter set enumerating sum-square error between circuit and model (scalar)
A	Signal amplitude
$b_m,\ a_m$	Transfer function numerator and denominator coefficients of mth degree
C	Capacitance
C_{cb}, C_{eb}	Collector-base and emitter-base junction capacitances of the BJT
EE	Elementary effect as used in the Morris method of sensitivity analysis
f	Frequency in Hertz
f_0	Distance between bins of the DFT in Hertz
f_{s}	Sampling frequency
G	Conductance
$H(s)$	Transfer function at a given value of the Laplace variable s
I	Current
$i,\ i_{\mathrm{s}}$	Number of iterations, number of sub-iterations
I_{kf}, I_{kr}	Forward and reverse knee currents of the BJT

I_{se}, I_{sc}	Leakage saturation currents of the BJT
I_{s}	Saturation current of a *pn* junction
j	Imaginary number $j^2 = -1$
k	Boltzmann's constant
N	Ideality factor/emissivity coefficient, for the BJT sometimes split into forward N_{f} and reverse N_{r}
n	Discrete time or frequency index
N_{e}, N_{c}	Leakage coefficients of the BJT
N_{s}	Number of samples in a signal
q	The charge on an electron
R	Resistance
r	Number of trajectories used in sensitivity analysis
R_{b}, R_{e}, R_{c}	Base, emitter, and collector terminal resistances of the BJT
s	Laplace variable, $s = \sigma + j\omega$
$t,\ b$	Tone stack treble and bass potentiometer positions
T_{K}	Temperature in Kelvin
V	Voltage
V_{c}	Supply voltage, for single supplies typically referring to the collector of a BJT but here also used for the emitter
V_{i}	Voltage input to a circuit or circuit model, excitation signal
V_{o}	Output voltage of a circuit or circuit model
V_{p}	Peak voltage, or scaling value to achieve specified peak voltage
V_{t}	Thermal voltage of a *pn* junction
$w(\cdot)$	Window function

Chapter 1

Introduction

"The numbers all go to eleven. Look, right across the board, eleven, eleven, eleven and..."

Nigel Tufnell, "This is Spinal Tap"

Analogue audio effects are devices which attract passionate views and loyalty from their users, perhaps due to the strong emotional connection a musician forms with their equipment when writing and playing music. With the development of digital technology and digital signal processing (DSP) reaching a point where such audio effects using DSP can be produced at scale, a competition has begun between analogue and digital effects. Musicians accustomed to the sound of analogue equipment have oft stated that the new digital effects are lacking of character or 'warmth' (e.g. [1]). This in turn has inspired the development of the field of Virtual Analogue (VA) which aims to capture the sound of analogue effects to a degree at which these musicians are convinced that no detail has been lost [2].

Research into VA has produced models of a wide variety of different circuits and a similarly wide range of techniques. One way of dividing the approaches to VA modelling is through determining how much information about the circuit is supposed *a priori*. This divides VA models into a sliding scale of classifications from white box to black box, white box modelling referring to a model built from physical first principles, whereas black box modelling uses only the response of a device to excitation, assuming only the bare minimum about the device's underlying behaviour [3]. Approaches at both extremes of the scale can successfully model the nonlinear behaviour that give many popular effects their signature sound (e.g. overdrive, distortion), though each has their own benefits.

White box models – referred to in this work as physical models due to their formation from physical phenomena – are designed using information extracted from

schematics. Schematic information is broadly divided into topology and a circuit's component parameter values, e.g. resistance. Perhaps the most notable example of physical modelling is SPICE - Simulation Programme with Integrated Circuit Emphasis, pioneered at UC Berkeley by Nagel and Pederson [4]. While SPICE software can simulate a diverse range of circuits and perform many useful analyses, fundamentally SPICE is intended for circuit designers and not musicians. In particular the feature missing from SPICE but demanded by digital audio effects is the ability to simulate in real-time, pushing the field towards algorithms described as "real-time SPICE" by David Yeh [5]. Such algorithms include Wave Digital Filters, Port Hamiltonian Systems, and state-space models, discussed in Chapter 2.

By designing these models from a precise and complete set of information given by our understanding of electronic components and how they perform when assembled in circuits results in models that inherently capture much of the behaviour of real circuits. However, several potential discrepancies exist: firstly should the information be taken from a schematic it does not refer to a real circuit. For example, component parameter values will be given within a tolerance and not to exact values. Secondly there may be components present that exhibit behaviour not sufficiently captured with existing models. Most authors in the field of VA treat this discrepancy as inconsequential, comparing their models to SPICE, e.g. [6, 7]. This does not account for the potential difference between SPICE and measurements of a specific device, which given the search for the unique qualities of analogue audio effects may be larger than assumed.

One method of designing a physical model from measurements of a real circuit is to measure each component individually. This process is arduous: while ubiquitous tools such as LCR meters and multimeters have been designed to measure the primary parameters of components such as resistors and capacitors, the measurements are difficult to perform while the components are connected in a circuit. To acquire an accurate direct measurement would necessitate the disassembly the circuit through desoldering. Further, components with more than two terminals typically require specialist equipment and software packages to extract parameter values suitable for simulation, for instance the Bipolar Junction Transistor [8].

Figure 1.1 shows an example diagram of a guitar pedal, labelling the elements that are exposed to the user as well as the internal electronic components. No knowledge of the internal circuit is required in the design of black box models, bypassing the need to measure each component. The model reproduces the behaviour of the circuit deriving the required information from measurements of the circuit's output in response to a given input.

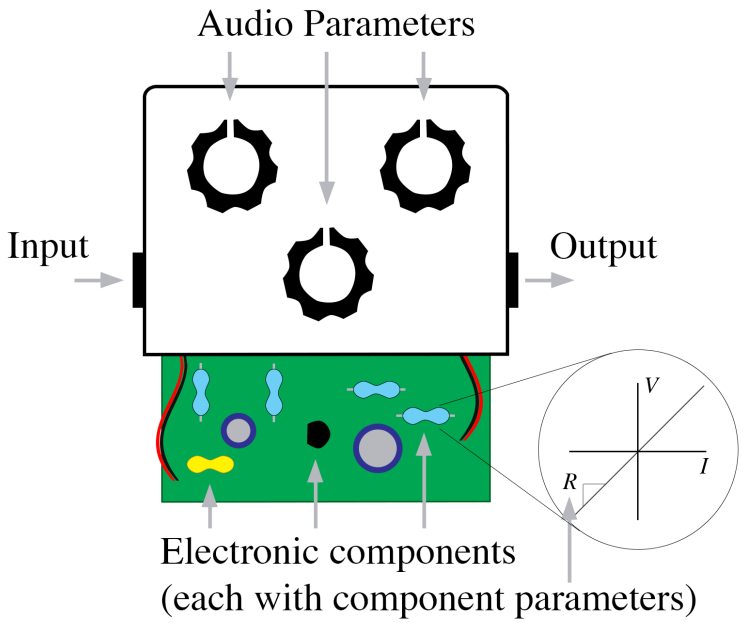


Figure 1.1: Cutaway diagram showing the internals of an guitar pedal, noting the difference between audio parameters – those exposed to the user in this case by knobs – and component parameters, which define the behaviour of a specific component like resistance (here noted by R).

This work utilises objective metrics to determine the fit of a circuit to its model, assuming that should the input/output behaviour be matched to a high degree across the anticipated range of input signals, that the model and circuit will be perceptually indistinguishable. Perceptual metrics exist for evaluating audio quality [9], but so far only an initial foray into using perceptual information for identification has been performed [10]. Further discussion around identification metrics is given in Section 2.2.3.

In addition to the input/output relation are the 'audio parameters', e.g. the knobs on the top of a pedal that a player tunes to produce their desired tone. Modelling potentiometers and their encompassing circuit physically automatically encapsulates the change in behaviour relative to a change in audio parameters. Within physical VA modelling literature exists specific research into how to update a model's audio parameters with efficiency, noting the importance for emulation [11]. In contrast, modelling through input/output measurements usually fixes the position of each audio parameter in one position to create the model (e.g. [12, 13]). Capturing multiple potentiometer positions therefore requires multiple identifications of the circuit and effectively multiple models based on each set of measurements.

1.1 Thesis overview

The main aim of this work is to identify guitar effect circuits using physical models and input/output measurements. This strategy establishes a link between the philosophies of black box and physical modelling, simultaneously avoiding the need of circuit disassembly while utilising the circuit's topology to capture changes in behaviour due to audio parameters.

The primary objectives of this work are founded in providing suitable models for use in the circuit identification, followed by the goals of said circuit identification:

- investigate and develop suitable root-finding algorithms for the efficient and robust simulation of nonlinear circuits using physical models;
- in the context of VA, analyse the suitability of existing component models for the germanium BJT - a component found in a selected case study for the identification;
- develop methods for minimising the difference in input/output behaviour between circuit and model through identification;
- develop methods for retrieving accurate values of component parameters from input/output measurements of a circuit, again through identification.

Both identification methods fundamentally require a circuit topology and set of initial parameters from which the identification procedure continues. A notable difference is seen between the two identification objectives. While the accurate estimation of physical component parameter values may initially seem like the clear objective when identifying circuits with physical models, it is not necessary in all modelling scenarios. A more utilitarian objective is simply the capture of the input/output behaviour of the circuit resulting in minimal observable differences, here noted by the term 'calibration'. Differences between the two identification objectives are further expounded in Chapter 5.

A complete set of contributions can be found in Section 8.1, but can be largely summarised as

- Two root-finding algorithms that utilise the form of the equation as given by the chosen modelling paradigm, offering improved robustness and efficiency in the presented case studies,

- Physical parameters as extracted from measurements of germanium BJTs. A comparison of the corresponding BJT models of different levels of complexity in the application of VA circuits, providing insight into when more complex models are suitable.

- A calibration strategy that improves the fit of a physical circuit model's output to that of the measured circuit. Optimisation time is drastically reduced by reducing search space dimensions through parameter screening.

- Redundancy in input/output models of circuits is detected and alleviated through the inclusion of a known component parameter, and is numerically demonstrated to retrieve circuit parameters from simulated measurements

- A parameter estimation strategy is presented which directly estimates each of the component parameters of a circuit using input/output measurements of a circuit, demonstrated in application to both a linear and nonlinear circuit.

Broadly the thesis progresses towards the circuit-level identification procedures first through a study of the background of the field, followed by studies focussed on elements that may prevent a successful identification, namely root-finding algorithms and the identification of the germanium BJT.

Chapter 2 presents a study of the surrounding literature with two areas of emphasis. The first is on the physical modelling of audio circuits, noting different modelling paradigms with specific detail into deriving models used in the remainder of the thesis. Models of individual components are also discussed and how more complex component models may be derived from measurements. The second focus is the exploration of existing identification strategies as applied to audio circuits, with specific strategies from each of black and grey box, and physical models. Metrics for comparing model and circuit are discussed, informing choices made in the identification strategy later discussed in Chapter 5.

Chapter 3 presents a comparative study into root-finding algorithms. Although this study is isolated from the main theme of identification, it is essential to have a robust and efficient root-finding algorithm to ensure that the identification procedure is successful. A range of relevant algorithms are discussed, and two new algorithms are presented that attempt to improve performance by using information derived from the circuit model. The two new algorithms are then compared with three existing algorithms to determine whether they offer improvement, and which algorithms are most suitable for the simulation of nonlinear circuit models.

The component-level identification of the germanium BJT is presented in Chapter 4. Through comparison of intial measurements of two germanium BJTs and one silicon BJT, a marked difference is found between semiconductor materials demonstrating the need for additional model complexity. The previously applied Ebers-Moll model is then extended towards the Gummel-Poon model, both of which are identified with a hybrid parameter extraction/optimisation approach. Finally, the identified models are used in circuit models to determine the difference between the BJT models of varying complexity when applied in a VA model.

Having tackled the pre-requisite issues required for the identification procedures, it is then possible to begin the discussion of circuit-level identification. An initial overview of the identification procedures is supplied in Chapter 5, with each element of both procedures progressively introduced for analysis resulting in a clear picture of their application to circuit measurements.

Calibration is discussed in full in Chapter 6. Using the results of the BJT identification, the Dallas Rangemaster Treble Booster is identified from input/output measurements of a real-world circuit. Through analysis of the identification problem, it is found that changes in some component parameter values results in little change in the output of the circuit, suggesting they may be fixed and therefore removed from the identification. Using this information, it is then investigated whether the optimisation becomes more efficient when operating on a reduced set of parameters. A further study of silicon vs. germanium BJTs is presented by exchanging the BJT of the Rangemaster, indicating which circuit identification produces a more accurate model.

Chapter 7 presents the second identification procedure: parameter estimation, from which component parameter values are the objective. Circuit models are analysed to determine the feasibility of estimating their component parameters, and a strategy of ensuring this property is found through the inclusion of a known component. A linear circuit with multiple audio parameters and a nonlinear circuit are both used as case studies.

Finally the thesis is concluded in Chapter 8, where the contributions of the thesis are summarised as well as the thesis on the whole. Future research directions are discussed for those inclined to pursue them.

Additional resources for the thesis are noted throughout the work in relevant places and hosted online[1].

[1] https://bholmesqub.github.io/thesis/

1.2 List of Publications

The following papers have been published over the course of researching the work in this thesis:

1. B. Holmes and M. van Walstijn. 'Improving the robustness of the iterative solver in state-space modelling of guitar distortion circuitry', Proceedings of the 18th International Conference on Digital Audio Effects, Trondheim, Norway, Dec. 2015.

2. B. Holmes and M. van Walstijn. 'Physical model parameter optimisation for calibrated emulation of the Dallas Rangemaster Treble Booster guitar pedal', Proceedings of the 19th International Conference on Digital Audio Effects, Brno, Czech Republic, Sept. 2016

3. B. Holmes, M. Holters, M. van Walstijn. 'Comparison of Germanium Bipolar Junction Transistor Models for Real-Time Circuit Simulation', Proceedings of the 20th International Conference on Digital Audio Effects, Edinburgh, Scotland, Sept. 2017

Chapter 2

Circuit modelling and identification: background

To frame the research within this thesis the preceding (and concurrent) work in physical modelling and identification of audio circuits must be expounded. The chapter is divided into two main sections: the first section discussing the derivation of the chosen physical modelling algorithm and alternate algorithms. The latter section then discusses circuit identification in VA and the complementary black and grey box models that have been utilised in the literature.

2.1 Physical circuit modelling algorithms

Many competing approaches of physical circuit modelling exist, each with different combinations of desirable properties such as stability, passivity, capability to handle nonlinear components etc. At the beginning of this research state space models were the most mature paradigm that could handle multiple nonlinearities, having already been used to model a common-emitter amplifier [14], the Fuzz Face [15], the MXR Phase 90 pedal [16], and the Dunlop Crybaby wah pedal [11]. As such, state space models were adopted for use in the circuit identification. Over the duration of this work, the gap of capabilities between state space modelling and other approaches has narrowed. Those approaching the problem today may have similar success with other modelling algorithms.

This section describes each of the competing approaches: Wave Digital Filters (WDF), Port Hamiltonian Systems (PHS), and state space models. Component models are first discussed, being largely independent from each approach and necessary for their construction. Focus is then given to the derivation of both state space models and

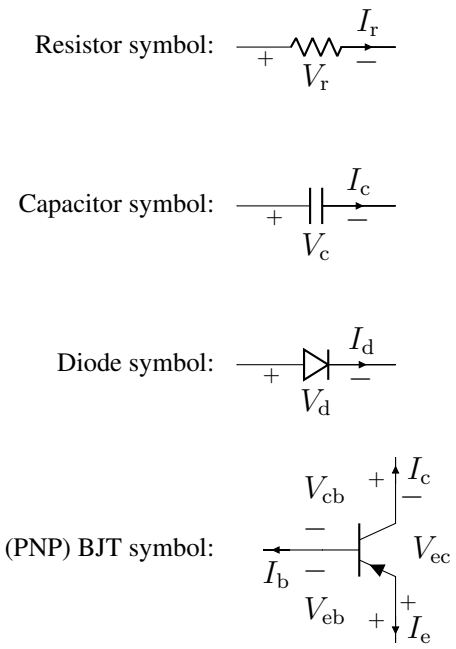


Figure 2.1: Schematic symbols for each of the components used throughout the thesis.

the other modelling paradigms used in the identification strategies, all of which are derived from the Modified Nodal Analysis (MNA) representation of a circuit.

2.1.1 Component modelling

In physical circuit modelling components are typically independent of the selected circuit-level modelling algorithm. Excluded from this assumption are more complex nonlinear components which often require additional considerations to enable compatibility (e.g. operational amplifiers in models derived using MNA [17]). Component models can be defined in terms of an *I-V*, or current-voltage, relationship, though auxiliary variables are common for describing the underlying physical process that links voltage to current. One or more component parameters are used in the definition of the *I-V* relationship, the value(s) of which is related to a specific instance of the component e.g. a resistor has an inherent *resistance*. Within this section the most commonly used component models are discussed and their parameters noted. Component parameters are of particular interest as they become the variables upon which the identification process operates.

A component model can be broadly defined by two properties: whether it is linear or nonlinear, and whether it is static or dynamic. The term 'dynamic' notes memory in the device, i.e. that the behaviour of the component is dependent upon past behaviour,

contrary to 'static' which indicates that the component's response is exclusively instantaneous. An example of a static component is the resistor, and for dynamic the capacitor. To create a model for simulation, dynamic components require discretisation in the time domain which will be discussed during the derivation of a computable model. This section focusses on the continuous time domain representations.

'Linearity' is a well defined mathematical concept, requiring a given function (in this case the *I-V* relation) to satisfy the two properties

- additivity: $f(x+y) = f(x) + f(y)$;
- homogeneity : $f(\alpha x) = \alpha f(x)$.

Simply put, nonlinear component models are any which do not satisfy these properties, for example, the diode and BJT.

In reality, every component will have some element of dynamic and nonlinear behaviour but component models frequently neglect these factors. This is to first place the focus on the *ideal* underlying behaviour as opposed to the parasitic/non-ideal. Further, the effect of non-ideal behaviour is often negligible. The component models discussed in this section are found from their fundamental underlying physical behaviour. Each case study used throughout the thesis consists of only four components (and ideal voltage sources): the resistor, capacitor, diode, and BJT.

Resistor

A resistor is defined by Ohm's law, parameterised using resistance R. Ohm's law can be equivalently represented by a conductance noted here by $1/R = G_{\mathrm{R}}$:

$$V_{\mathrm{r}} = I_{\mathrm{r}} R, \quad I_{\mathrm{r}} = V_{\mathrm{r}} G_{\mathrm{R}}. \tag{2.1}$$

Ohm's law defines the voltage as proportional to the current, the specified proportion being the resistance.

Capacitor

From a physical perspective of the capacitor, voltage is proportional to the charge Q on the capacitor, $V = Q/C$. As with the resistor the proportionality is controlled by a parameter, here capacitance C. Charge is equal to the integral of the current with respect to time, leading to the definition of the *I-V* relationship,

$$V_{\mathrm{c}} = \frac{1}{C} \int I_{\mathrm{c}} \, dt, \quad I_{\mathrm{c}} = C \frac{dV_{\mathrm{c}}}{dt}. \tag{2.2}$$

Capacitor memory is clear by the presence of the integral/derivative with respect to time. The model is linear as one quantity is proportional to the other, only shifted in time.

The inductor is a counterpart of the capacitor, where the current is proportional to the integral of the voltage. It is not featured in any circuits studies throughout this work and therefore its mathematical definition and any further discussion about it is omitted.

Diode

A frequently used model of the diode is found in the Shockley model [18, 19, 20]:

$$I_{\mathrm{d}} = I_{\mathrm{s}} \left(\mathrm{e}^{\frac{V_{\mathrm{d}}}{N V_{\mathrm{t}}}} - 1 \right). \tag{2.3}$$

Parameters of the model are saturation current, I_{s}, and ideality factor/emissivity coefficient N. The term V_{t} is the thermal voltage and is not a parameter, its value is defined by the temperature of the *pn*-junction, i.e.

$$V_{\mathrm{t}} = T_{\mathrm{K}} \frac{k}{q}, \tag{2.4}$$

where T_{K} is the temperature in Kelvin, k is Boltzmann's constant and q is the charge on an electron. The value of V_{t} is not inherent to a specific diode; it would be equivalent for multiple diodes so long as the temperature of the *pn*-junction is the same.

As the diode is nonlinear and cannot be represented by a linear impedance, the component is often modelled using a voltage-controlled current source (VCCS) that directly represents the *I-V* relationship (e.g. [18]). The same is true of the BJT which instead uses two VCCS connected.

Bipolar Junction Transistor

The most commonly used model for the BJT in VA is the Ebers-Moll model, e.g. [15, 11, 16]. This model can be expressed by 2 of 3 *I-V* relationships (the third inferred from superposition):

$$I_{\mathrm{b}} = \frac{I_{\mathrm{s}}}{\beta_{\mathrm{f}}} \left(\mathrm{e}^{\frac{V_{\mathrm{eb}}}{N V_{\mathrm{t}}}} - 1 \right) + \frac{I_{\mathrm{s}}}{\beta_{\mathrm{r}}} \left(\mathrm{e}^{\frac{V_{\mathrm{cb}}}{N V_{\mathrm{t}}}} - 1 \right), \tag{2.5}$$

$$I_{\mathrm{c}} = I_{\mathrm{s}} \left(\mathrm{e}^{\frac{V_{\mathrm{eb}}}{N V_{\mathrm{t}}}} - 1 \right) - I_{\mathrm{s}} \frac{\beta_{\mathrm{r}} + 1}{\beta_{\mathrm{r}}} \left(\mathrm{e}^{\frac{V_{\mathrm{cb}}}{N V_{\mathrm{t}}}} - 1 \right). \tag{2.6}$$

Note that $V_{\mathrm{ec}} = V_{\mathrm{eb}} - V_{\mathrm{cb}}$, and $I_{\mathrm{e}} = -(I_{\mathrm{c}} + I_{\mathrm{b}})$. The parameters of the Ebers-Moll model are I_{s} and N, directly from the Shockley diode model, and β_{f} and β_{r} which are the forward and reverse current gains respectively.

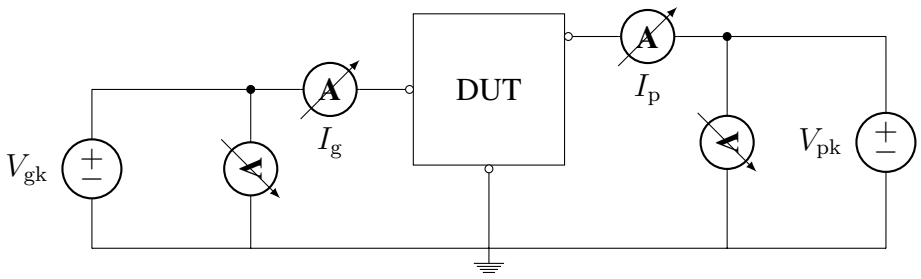


Figure 2.2: Measurement setup for simultaneously capturing voltage and current data of a vacuum tube [28].

Additional nonlinear component models

Circuits favoured by analogue-enthusiasts utilise a wide variety of components that exhibit unique nonlinear behaviour. Although limited to diodes and BJTs in this work, researchers in the field of VA have modelled many more components. Examples of this include the nonlinear behaviour of the op-amp as the output approaches the power supply voltage [17], and the saturating behaviour of the transformer [21, 22].

One component that has been of particular focus in the literature is the valve or vacuum tube, used in [23, 7, 24] among others. Recent developments include a phenomenological model presented by Koren in [25], a physical-interpolative approach by Cardarilli [26], and a physically motivated model by Dempwolf [27], which while still using free, non-physical parameters, overcomes the an issue present in the other two models: discontinuities in the *I-V* derivatives.

Studies of the vacuum tube provide important context to this work as many are based on measurements of a real device. The tube is a 3-terminal, nonlinear device and as such demands a similar measurement strategy to the BJT. Measurements of the tubes are a necessity because the models are novel and not simply adapted from existing models to work in real-time modelling paradigms. Equipment is required to simultaneously measure the grid and plate current, and grid-cathode and plate-cathode voltages of the device under test (DUT). One example of such a measurement setup was utilised in [28], and is repeated in Figure 2.2. Multiple surfaces are found, e.g. V_{pk}, V_{gk} vs. I_{p}, from which component model parameters are extracted, in [29] using curve-fitting algorithms. The result is a component model that directly relates to a given device, and should the device be utilised in a circuit that is being modelled, having exact parameter values will yield a circuit model with high accuracy, as validated in [29] in the case of a common-cathode amplifier.

2.1.2 Modified Nodal Analysis

To derive each circuit model in this work, MNA is used. MNA has been used frequently in model derivations in recent literature, most notably [30, 14, 11, 31]. The original publication by Ho [32] presented the work to address the limitations of traditional nodal analysis, the most important of which for VA circuit models being the inability to model voltage sources.

Nodal analysis expresses a circuit as a set of nodes, utilising Kirchoff's Current Law (KCL) which states that the sum of the currents entering and exiting a node is equal to zero:

$$0 = \sum_{k=1}^{N} I_k. \tag{2.7}$$

Used in combination with a component's *I-V* relation, for linear components with only current sources, a circuit is generally expressed as

$$\mathbf{G}\mathbf{v} = \mathbf{i}, \tag{2.8}$$

where $\mathbf{G}$ contains the circuit conductances placed to link nodal voltages and current sources. Within this work, resistors and capacitors are the only linear component models used, for which $\mathbf{G}$ can be decomposed into

$$\mathbf{G} = \mathbf{N}_{\mathrm{R}}^{\mathrm{T}}\mathbf{G}_{\mathrm{R}}\mathbf{N}_{\mathrm{R}} + \mathbf{N}_{\mathrm{C}}^{\mathrm{T}}\mathbf{G}_{\mathrm{C}}\mathbf{N}_{\mathrm{C}}. \tag{2.9}$$

Here $\mathbf{G}_{\mathrm{R}}$ and $\mathbf{G}_{\mathrm{C}}$ are diagonal matrices containing the resistances and capacitances. The connections of the components are restricted to the incidence matrices $\mathbf{N}_{\mathrm{R}}$ and $\mathbf{N}_{\mathrm{C}}$, which contain values 1 and -1 to indicate connections.

To adapt nodal analysis to enable the use of voltage sources, the model is extended to include $\mathbf{u}$ which contains the known voltage sources, and $\mathbf{i}_{\mathrm{v}}$ which are the respective unknown currents through the voltage sources. The admittance matrix, $\mathbf{G}$, is then extended to contain the connections of the voltage sources, resulting in

$$\mathbf{S} = \begin{bmatrix} \mathbf{G} & \mathbf{N}_{\mathrm{U}} \\ \mathbf{N}_{\mathrm{U}} & \mathbf{0} \end{bmatrix}, \tag{2.10}$$

where $\mathbf{N}_{\mathrm{U}}$ is an incidence matrix specifying the connections of the voltage sources. The zero-value elements in the lower right of $\mathbf{S}$ will remain empty for each of the models in this work, though are required for circuits featuring components such as op-amps, modelled with voltage controlled voltage sources [17]. The full MNA model is expressed by

$$\mathbf{S}\begin{bmatrix} \mathbf{v} \\ \mathbf{i}_{\mathrm{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{i} \\ \mathbf{u} \end{bmatrix}. \tag{2.11}$$

With this form each of the chosen case studies in this work can be modelled. Voltage controlled current sources in $\mathbf{i}$ can represent many nonlinear component models like the BJT. The full MNA form as given in (2.11) is inefficient however as even linear circuits require a matrix inversion at each time step that solves for each of the unknown voltages. This expense can be reduced as it is uncommon to require the voltage at each node. For linear circuits, a model that condenses the input/output behaviour into a few calculations can be found in transfer functions. As previously mentioned for nonlinear circuits state space models will be used, as derived through the Nodal DK method.

Transfer functions

Transfer functions model linear circuits in the frequency domain as a relation between a single input and output. Commonly voltages are the quantity of interest, used to drive circuits and simultaneously measured at the output, resulting in a transfer function given by

$$H(s) = \frac{V_{\mathrm{o}}(s)}{V_{\mathrm{i}}(s)} = \frac{b_m s^m + b_{m-1}s^{m-1} \ \ldots \ + b_1 s + b_0}{a_m s^m + a_{m-1}s^{m-1} \ \ldots \ + a_1 s + a_0}. \tag{2.12}$$

Here s is the Laplace variable and b_m and a_m are the coefficients that define the function's response. Though, typically, the Laplace variable refers to the sum of a steady state term $j\omega$ and transient term σ i.e. $s = \sigma + j\omega$, the signals of interest are assumed to be steady-state. In the steady-state case, $\sigma = 0$, which reduces the Laplace transform to the Fourier transform, though the Laplace variable is used to remain consistent with literature.

In some cases, it is useful to find linear models of circuits exhibiting nonlinear behaviour, for example to provide a course approximation of the circuit's frequency response. Nonlinear component models must be linearised to be used in a transfer function. A linearised component model is found by taking the derivative of the nonlinear function at an operating point. This model can then be used as a conductance in $\mathbf{G}$. In the instance of a diode for example, the diode's linear conductance is given by $G_{\mathrm{D}} = \frac{I_{\mathrm{s}}}{NV_{\mathrm{t}}} \mathrm{e}^{\frac{V_{\mathrm{q}}}{NV_{\mathrm{t}}}}$ where V_{q} is the quiescent voltage or voltage at the selected operating point.

Integration with respect to time in the time domain becomes $1/s$ in the frequency domain, so the frequency domain component model of the capacitor becomes $V(s) = I(s)/(sC)$.

One transfer function describes the relationship between one input and one output of a system. First the input must be selected, with remaining voltage inputs becoming open circuits which may require some adaptation to ensure there are no hanging nodes

(nodes with only one connection that cause computational issues). The vector of nodal voltages and currents through voltage sources is then solved for by inverting $\mathbf{S}$,

$$\begin{bmatrix} \mathbf{v} \\ \mathbf{i}_{\mathrm{v}} \end{bmatrix} = \mathbf{S}^{-1} \begin{bmatrix} \mathbf{i} \\ \mathbf{u} \end{bmatrix}. \tag{2.13}$$

An additional incidence vector must then be created to select the output, $\mathbf{N}_{\mathrm{O}}$, which contains 1 at the positive node and -1 at the negative node. If the output is referenced to ground then only one of the non-zero values is required with polarity at the discretion of the modeller. Multiplying both sides of (2.13) by $\mathbf{N}_{\mathrm{O}}$ equates the system to the output voltage V_{o},

$$V_{\mathrm{o}} = \mathbf{N}_{\mathrm{O}} \begin{bmatrix} \mathbf{v} \\ \mathbf{i}_{\mathrm{v}} \end{bmatrix} = \mathbf{N}_{\mathrm{O}}\, \mathbf{S}^{-1} \begin{bmatrix} \mathbf{i} \\ \mathbf{u} \end{bmatrix}. \tag{2.14}$$

The result is a scalar equation from which the ratio $V_{\mathrm{o}}/V_{\mathrm{i}}$ can be found, i.e. the transfer function.

Nodal DK method

The Nodal DK method uses a state space model that is capable of modelling nonlinear behaviour in circuits. This method originated as the K-method by Borin et. al. [33], named after Kirchoff variables to distinguish it from WDFs. The K-method performs transformations on time-dependent and nonlinear equations to alleviate delay-free loops, enabling solutions to implicit equations to be stored in a LUT and thus preventing online root-finding. This was then adapted by Yeh to directly discretise dynamic components use and MNA to enable algorithmic derivation of circuit models. A discussion of the method refers to it as the Nodal DK method [11], and it is from this source that the description in this section is informed.

Dynamic components are discretised directly using the trapezoidal rule. For the capacitor:

$$\frac{1}{2}\left(I_{\mathrm{C}}(n) + I_{\mathrm{C}}(n-1)\right) = \frac{C}{T}\left(V_{\mathrm{C}}(n) - V_{\mathrm{C}}(n-1)\right), \tag{2.15}$$

where T is the sampling period. As with previous component models the desire is to find an *I-V* relationship, which can be found through expressing the discretised capacitor as

$$I_{\mathrm{C}}(n) = G_{\mathrm{C}}\left(V_{\mathrm{C}}(n) - V_{\mathrm{C}}(n-1)\right) - I_{\mathrm{C}}(n-1), \tag{2.16}$$

where $G_{\mathrm{C}} = 2C/T$ for the Nodal DK method. The two historic quantities can be grouped into a state term, such that

$$I_{\mathrm{C}}(n) = G_{\mathrm{C}} V_{\mathrm{C}}(n) - x(n-1), \tag{2.17}$$

where

$$x(n) = G_{\mathrm{C}} V_{\mathrm{C}}(n) + I_{\mathrm{C}}(n). \tag{2.18}$$

To find the recurrence relation that updates the state variable x, (2.17) is substituted into (2.18) to find

$$x(n) = 2G_{\mathrm{C}} V_{\mathrm{C}}(n) - x(n-1). \tag{2.19}$$

To derive this form from the full MNA, the RHS vector from (2.11) can be decomposed into the relevant quantities,

$$\mathbf{S} \begin{bmatrix} \mathbf{v} \\ \mathbf{i}_{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{N}_{\mathrm{C}} \\ \mathbf{0} \end{bmatrix} \mathbf{x}(n-1) + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{u}(n) + \begin{bmatrix} \mathbf{N}_{\mathrm{N}} \\ \mathbf{0} \end{bmatrix} \mathbf{f}(\mathbf{v}(n)), \tag{2.20}$$

which shows that the previous state $\mathbf{x}(n-1)$ forms a current source where the capacitors are placed, determined by $\mathbf{N}_{\mathrm{C}}$. This is also true for the nonlinear components with $\mathbf{f}(\mathbf{v}(n))$ and $\mathbf{N}_{\mathrm{N}}$. The system matrix $\mathbf{S}$ is inverted to find the solution for the LHS vector,

$$\begin{bmatrix} \mathbf{v} \\ \mathbf{i}_{\mathbf{v}} \end{bmatrix} = \mathbf{S}^{-1} \left(\begin{bmatrix} \mathbf{N}_{\mathrm{C}} \\ \mathbf{0} \end{bmatrix} \mathbf{x}(n-1) + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{u}(n) + \begin{bmatrix} \mathbf{N}_{\mathrm{N}} \\ \mathbf{0} \end{bmatrix} \mathbf{f}(\mathbf{v}(n)) \right). \tag{2.21}$$

Multiplying the equation by the relevant incidence matrices reveals the state space matrices for each variable, e.g. multiplying by $\mathbf{N}_{\mathrm{N}}$ produces $\mathbf{D}_{\mathrm{N}} - \mathbf{F}_{\mathrm{N}}$. An exception to this is for the state update, which requires the additional terms from (2.19), i.e.

$$\mathbf{x}(n) = 2\, \mathbf{G}_{\mathrm{C}} \mathbf{S}^{-1} \left(\begin{bmatrix} \mathbf{N}_{\mathrm{C}} \\ \mathbf{0} \end{bmatrix} \mathbf{x}(n-1) + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{u}(n) + \begin{bmatrix} \mathbf{N}_{\mathrm{N}} \\ \mathbf{0} \end{bmatrix} \mathbf{f}(\mathbf{v}(n)) \right) - \mathbf{x}(n-1). \tag{2.22}$$

The complete discrete-time state space model is given by the state update, output, and nonlinear function equations,

$$\mathbf{x}(n) = \mathbf{A}\mathbf{x}(n-1) + \mathbf{B}\mathbf{u}(n) + \mathbf{C}\mathbf{f}(\mathbf{v}_{\mathrm{n}}(n)), \tag{2.23}$$

$$\mathbf{y}(n) = \mathbf{D}_{\mathrm{O}}\mathbf{x}(n-1) + \mathbf{E}_{\mathrm{O}}\mathbf{u}(n) + \mathbf{F}_{\mathrm{O}}\mathbf{f}(\mathbf{v}_{\mathrm{n}}(n)), \tag{2.24}$$

$$\mathbf{v}_{\mathrm{n}}(n) = \mathbf{D}_{\mathrm{N}}\mathbf{x}(n-1) + \mathbf{E}_{\mathrm{N}}\mathbf{u}(n) + \mathbf{F}_{\mathrm{N}}\mathbf{f}(\mathbf{v}_{\mathrm{n}}(n)), \tag{2.25}$$

where $\mathbf{x}$ is the model state, $\mathbf{y}$ the output, and $\mathbf{u}$ the input. Matrices $\mathbf{A} - \mathbf{C}$, $\mathbf{D}_{\mathrm{O}} - \mathbf{F}_{\mathrm{O}}$ and $\mathbf{D}_{\mathrm{N}} - \mathbf{F}_{\mathrm{N}}$ control the weighting of each variable used in the update of the model for state, output, and nonlinearities respectively. The generic nonlinear function $\mathbf{f}(\mathbf{v}_{\mathrm{n}}(n))$

contains the relevant *I-V* relationships. The state space matrices are given by

$$\mathbf{A} = 2\mathbf{G}_{\mathrm{C}} \begin{bmatrix} \mathbf{N}_{\mathrm{C}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{N}_{\mathrm{C}} & \mathbf{0} \end{bmatrix}^{\mathrm{T}} - \mathbf{I}, \tag{2.26}$$

$$\mathbf{B} = 2\mathbf{G}_{\mathrm{C}} \begin{bmatrix} \mathbf{N}_{\mathrm{C}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix}^{\mathrm{T}}, \tag{2.27}$$

$$\mathbf{C} = 2\mathbf{G}_{\mathrm{C}} \begin{bmatrix} \mathbf{N}_{\mathrm{C}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{N}_{\mathrm{N}} & \mathbf{0} \end{bmatrix}^{\mathrm{T}}, \tag{2.28}$$

$$\mathbf{D}_{\mathrm{O}} = \begin{bmatrix} \mathbf{N}_{\mathrm{O}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{N}_{\mathrm{C}} & \mathbf{0} \end{bmatrix}^{\mathrm{T}}, \tag{2.29}$$

$$\mathbf{E}_{\mathrm{O}} = \begin{bmatrix} \mathbf{N}_{\mathrm{O}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix}^{\mathrm{T}}, \tag{2.30}$$

$$\mathbf{F}_{\mathrm{O}} = \begin{bmatrix} \mathbf{N}_{\mathrm{O}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{N}_{\mathrm{N}} & \mathbf{0} \end{bmatrix}^{\mathrm{T}}, \tag{2.31}$$

$$\mathbf{D}_{\mathrm{N}} = \begin{bmatrix} \mathbf{N}_{\mathrm{N}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{N}_{\mathrm{C}} & \mathbf{0} \end{bmatrix}^{\mathrm{T}}, \tag{2.32}$$

$$\mathbf{E}_{\mathrm{N}} = \begin{bmatrix} \mathbf{N}_{\mathrm{N}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix}^{\mathrm{T}}, \tag{2.33}$$

$$\mathbf{F}_{\mathrm{N}} = \begin{bmatrix} \mathbf{N}_{\mathrm{N}} & \mathbf{0} \end{bmatrix} \mathbf{S}^{-1} \begin{bmatrix} \mathbf{N}_{\mathrm{N}} & \mathbf{0} \end{bmatrix}^{\mathrm{T}}. \tag{2.34}$$

To calculate one time step of the model, first the nonlinear equation (2.25) must be solved, which can be phrased as a root-finding problem by introducing a new function

$$\mathbf{g}(\mathbf{v}_{\mathrm{n}}(n)) = \mathbf{p}(n) + \mathbf{F}_{\mathrm{N}}\mathbf{f}(\mathbf{v}_{\mathrm{n}}(n)) - \mathbf{v}_{\mathrm{n}}(n) = \mathbf{0}, \tag{2.35}$$

where $\mathbf{p}(n) = \mathbf{D}_{\mathrm{N}}\mathbf{x}(n-1) + \mathbf{E}_{\mathrm{N}}\mathbf{u}(n)$. The value of $\mathbf{v}_{\mathrm{n}}$ that satisfies this equation is then used in calculating the output at the current time step and updating the state variable.

Parameter focussed representation

One of the main objectives of this thesis is to estimate the values of component parameters through an identification process that uses measurements of a circuit's output in response to a given input. A factor that determines the feasibility of this objective is the parametric structure of the employed circuit model, i.e. how the parameters appear in the model relating input to output. To that end the MNA-derived models utilised in this work require analysis to determine whether their parameters can be directly estimated.

The selected analysis that is applied to case studies starting in Section 5.1.3 uses symbolic algebra as applied in ecological models in [34]. A common flaw is found in transfer functions derived from circuits: the transfer function coefficients contain

highly complex combinations of component parameters, as noted in e.g. [35, 36, 37]. A similar issue is present in the elements of state space model matrices. Complexity in the combinations of parameters used in the model increases the complexity of the analysis, increasing the difficulty in detecting whether the analysis has been performed correctly.

It therefore serves to find a similar model form to transfer functions and state space models with an equivalent parametric structure of these models but with reduced complexity. Three general simplifications can be achieved starting with the general MNA form shown in (2.11):

- Ideal voltage sources contribute no parameters towards the model and therefore can be removed and replaced with open circuits to preserve the circuit topology.
- Factors related to time or frequency such as those introduced by discretisation can be omitted. The model is not used for computation and therefore a continuous time domain model can be used.
- In the continuous time domain resistors operate on voltages $\mathbf{v}$, while capacitors operate on voltage derivatives $\mathbf{v}'$ meaning that combinations of resistors and capacitors can be isolated from each other.

Begin with a circuit model as defined by the general MNA form shown in (2.11). Voltage sources contribute $\mathbf{u}$ and corresponding currents $\mathbf{i}_\mathrm{v}$ in the model. Removing these vectors returns the MNA to the nodal form in (2.8) with one key difference: the topology including voltage sources has not been altered but the model has fewer elements, simplifying the anticipated analysis.

A transform to the frequency domain or discretisation is not required as the model's parametric structure is not affected by these processes. Resistors and capacitors must then be separated to act on nodal voltages and voltage derivatives independently, resulting in

$$\underline{\mathbf{N}_\mathrm{R}}^\mathrm{T}\,\underline{\mathbf{G}_\mathrm{R}}\,\underline{\mathbf{N}_\mathrm{R}}\,\mathbf{v} + \underline{\mathbf{N}_\mathrm{C}}^\mathrm{T}\,\underline{\mathbf{G}_\mathrm{C}}\,\underline{\mathbf{N}_\mathrm{C}}\,\mathbf{v}' = \underline{\mathbf{N}_\mathrm{N}}\,\underline{\mathbf{f}(\mathbf{v}_\mathrm{n})}. \tag{2.36}$$

The only current sources in $\mathbf{i}$ will be those used to model nonlinear components through VCCS, and so is instead expressed as $\mathbf{f}(\mathbf{v}_\mathrm{n})$. Underlined in blue are all matrices that relate to the topology of the modelled circuit, and in bold red are matrices that contain parameters. A complete parametric structure of the MNA form is contained within this model, with the only variables that are not component-parameters being the nodal voltages which are required to specify the circuit topology.

The application of this model form may not be immediately clear. Determining whether parameters can be estimated through symbolic algebraic analysis is applied to case studies beginning in Section 5.1.3, where this model form is first utilised. The parameter focussed representation will be shown (for the given case studies) to provide equivalent results but with a more visually intuitive format.

2.1.3 Circuit-level modelling paradigms

Wave Digital Filters

Linear Wave Digital Filters (WDFs) were introduced in [38] with a broad presentation of their uses in [39]. Electronic circuits are modelled in the wave domain in a modular approach where the passivity of the circuit is maintained. Kirchoff variables voltage and current are transformed into wave variables, incident and reflected waves a and b,

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & R_0 \\ 1 & -R_0 \end{bmatrix} \begin{bmatrix} V \\ I \end{bmatrix}, \tag{2.37}$$

where R_0 is the port impedance, unrelated from resistances in circuits as discussed thus far. Through careful selection of the value of R_0, one port is chosen to be reflection free which can be used to break a delay-free loop, and thus enables the inclusion of a single nonlinearity in a model, thoroughly discussed in [40] and exemplified in e.g. [19].

When addressing multiple nonlinearities the situation becomes more challenging. Perhaps the most direct method of breaking delay-free loops is the insertion of unit delay elements, used for example in the modelling of circuits featuring tubes [41, 42].

A second approach is found by iterating over the whole model in an artificial dimension to circumvent the issue of delay-free loops. This was proposed in [43] where it is applied to a bridged-T notch filter, which although linear has a challenging topology which requires a similar strategy as when facing a nonlinearity. The approach is further applied to a diode clipper in [44] with an improved iteration strategy, and for which further root finding algorithms are investigated in [45].

Most prevalent in recent literature is the use of R-type adapters. Use of graph theory from which R-type adapters originate was first applied to understand the underlying structure of WDFs in [46, 47]. Use of MNA was applied for the derivation of the R-type adapter in [31] with application for multiport components, multiple nonlinearities, and challenging topologies in [48, 49]. This was further applied in [7] with special concern applied to the root finding algorithms in [50].

The types of circuits that can be successfully modelled using WDFs and an R-type adapter has been further extended. Circuits featuring op-amps pose a challenge in that, instead of a VCCS as utilised in the modelling of diodes and BJTs, a voltage-controlled voltage source (VCVS) is required. A second challenge is posed by the connection of multiple VCCS elements, as in the case of series diodes. Forbidden WDF topologies for op-amps and series diodes were overcome in [51] by generalising the choice of the in/dependent variable placed at the root of the WDF structure.

Arbitrary topologies and active circuits were approached systematically in [52]. Four junctions were proposed for the unbounded WDF junction, as necessary for circuits such as the Baxandall tone control. Solutions were also proposed for handling the negative port resistance R_0 that occur in active circuits, localising the resultant complex valued waves that arise. These contributions facilitate the systematic modelling of circuits previously out of scope of WDF which previously required specific solutions.

Port-Hamiltonian Systems

Port-Hamiltonian Systems (PHS) are derived to conserve the total energy in a system, naturally capturing the power balance between circuit components. Port describes the approach of the decomposition of complex systems into a set of interconnected blocks. The approach has been applied successfully to several nonlinear audio case studies (both electronic and electro-acoustic) in the wah-pedal [53], a loudspeaker [54], and a Fender Rhodes piano [55]. An automated strategy of developing a PHS model with the use of a component model library was presented in [20], and given by

$$\left[\begin{array}{c} \frac{d\mathbf{x}}{dt} \\ \hline \mathbf{w} \\ \hline -\mathbf{y} \end{array}\right] = \left[\begin{array}{c|c|c} \mathbf{J_x} & -\mathbf{K} & -\mathbf{G_x} \\ \hline \mathbf{K}^{\mathrm{T}} & \mathbf{J_w} & -\mathbf{G_w} \\ \hline \mathbf{G_x}^{\mathrm{T}} & \mathbf{G_w}^{\mathrm{T}} & \mathbf{J_y} \end{array}\right] \left[\begin{array}{c} \nabla\mathcal{H}(\mathbf{x}) \\ \hline \mathbf{z}(\mathbf{w}) \\ \hline \mathbf{u} \end{array}\right] \tag{2.38}$$

where $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{u}$ are the same for state-space models. Component models are defined by their dissipated power with variable $\mathbf{w}$ and function $\mathbf{z}(\mathbf{w})$ which are related by $D = z(w)w$. For each electronic component $D = IV$, for example, in the case of the resistor let $w = I$ such that $z(w) = IR = V$.

Recent advances in PHS include direct inclusion of anti-aliasing into PHS models, achieved in [56] with the application of continuous time domain trajectories. Additionally each quantity used in the model has been successfully derived from the power balance [57].

State space models

Further developments have occurred in the state space modelling of circuits since the development of the Nodal DK method. Using the Nodal DK method as a base, in [11] the model was decomposed using the Woodbury identity to isolate a parametric section of the model, reducing the computational expense of updating the model when audio parameter values are changed.

A blockwise decomposition was applied in [58] where nonlinearities were separated into cascaded pairs. A higher number of nonlinear equations are solved overall but due to their division into smaller problems, the overall computational expense is lower. The decomposition also enabled approximations to be performed to reduce the number of nonlinear equations being solved concurrently. Improved accuracy is achieved in [59] where connections between sections are broken, but the connecting currents are found numerically.

Cases exist in which MNA is incapable of modelling a circuit, for example where multiple nonlinearities are connected directly. By application of a different underlying paradigm (i.e. not MNA), [60] alleviates these issues, producing a state space model with two series connected diodes without combining the nonlinearities into one function. This is further applied to a nonlinear model of a transformer using the Jiles-Atherton model in [22].

2.2 Circuit model identification

2.2.1 Input/Output measurement

Black box

A popular option in recent literature is the tuning of black box models using a swept sine measurement. A method utilising exponential swept sines was proposed in [62], designed for the identification of weakly nonlinear systems. An inverse filter was derived from the swept sine to deconvolve the input from the output in the time domain, resulting in a what could be described as a nonlinear impulse response.

This method has been used to identify many audio circuits with black box models, including using Chebyshev polynomials for an overdrive circuit [12], nonlinear convolution, Weiner models, and Hammerstein models on a dynamic range effect [63]. Recent revisions have improved the method to enable deconvolution in the frequency domain which is often more efficient, and the accurate capture of phase behaviour by

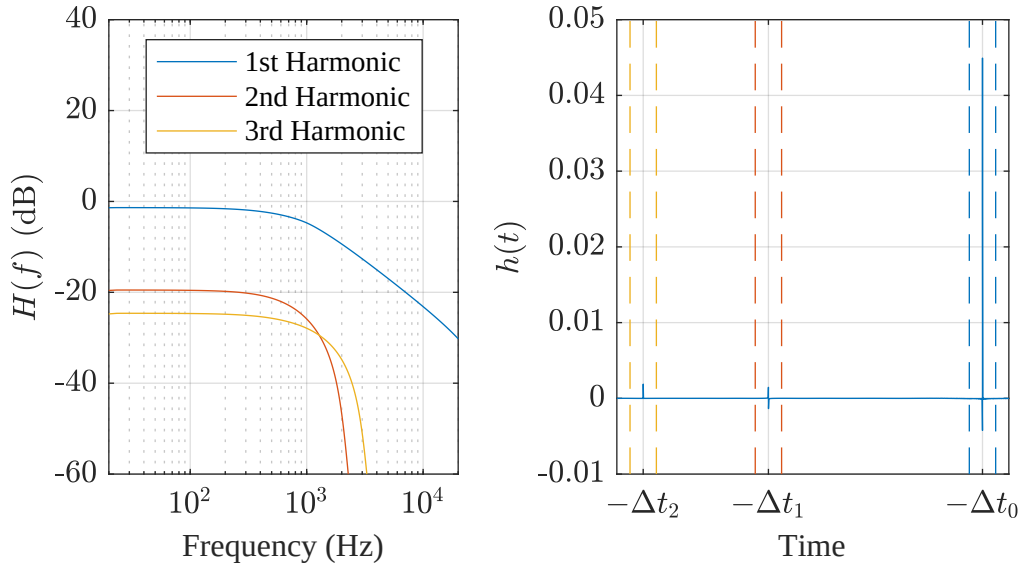


Figure 2.3: Nonlinear impulse response of the single sided diode clipper up to the 3rd harmonic. (left) Frequency domain amplitude response, (right) time domain impulses as separated by limits derived in [61].

a modification to the swept-sine signal to ensure that it is synchronised [61].

The synchronised swept-sine is defined as [61]

$$x(t) = \sin\left(2\pi f_1 L \mathrm{e}^{t/L}\right), \text{ where } L = \frac{1}{f_1}\mathrm{round}\left(\frac{\hat{T} f_1}{\ln(f_2/f_1)}\right). \tag{2.39}$$

Starting and end frequencies are noted by f_1 and f_2, and $\hat{T}$ is the approximate duration in seconds; the real duration modified by the function round which rounds the value to the nearest integer. To find the nonlinear impulse response an inverse filter of $x(t)$ is designed using the Fourier transform to efficiently deconvolve the input from the output in the frequency domain.

Figure 2.3 shows an example swept-sine analysis of a single sided diode clipper, introduced in Chapter 3. The right plot shows the time-domain impulse responses of the first 3 harmonics separated by windows marked by the dashed lines. The amplitude response of each of these windowed 3 harmonics is shown in the left hand plot. Despite not knowing the topology or physical behaviour of the circuit, the swept-sine has captured the low-pass behaviour of the circuit, as well as the asymmetry of the nonlinearity indicated by the presence of the 2nd harmonic.

The amplitudes of distortion harmonics are dependent upon the input amplitude of

the signal[1]. As the described swept-sine approach maintains a constant amplitude over the frequency range the resultant nonlinear impulse response will only be valid for that amplitude of input signal. For example, in the application of a diagonal Volterra series model if a different input amplitude is used, the corresponding output waveform show marked error [64].

To overcome changes in the nonlinear impulse response requires a method of adapting each response relative to the input signal's amplitude. A 'continuous' set of Volterra kernals were found in [64] through interpolation of kernels measured at set levels. To reduce the number of interpolation operations required, a single interpolated function is proposed in [65] that is shared across all kernels, but notes this will not likely be sufficient for strongly nonlinear devices. Results from both publications demonstrate accurate modelling of the case study circuits over the anticipated amplitude ranges.

Audio parameters present a similar problem as input amplitude to black-box models: should an audio parameter value change, the model's output will no longer necessarily fit that of the circuit. Only one publication has approached the subject, modelling the gain control on a guitar amplifier using Recurrent Neural Networks (RNN) [66]. The gain control is treated as an additional input to the RNN during training. Results of the RNN show a close match between the amplifier and model output, but only a single gain setting is used in the results preventing the confirmation of success at modelling audio parameters.

Grey box

In the case that some information about the audio circuit can be determined the identification problem is referred to as grey box. Within the field of VA the most notable example comes from industry in the form of Kemper amplifiers, a 'profiling' amplifier that identifies amplifier circuits using a grey box model [67].

This hardware has inspired a chain of research to surpass the level of fidelity achieved by Kemper, while also providing analysis and reproducable results. The method first utilised a Wiener-Hammerstein approach with a static nonlinearity and some additional elements implemented to mimic the circuit bias, sandwiched between two parametric equalizers [13]. Iterative optimisation was applied in several targeted

[1]This is visualised in a video on the accompanying website by changing the input amplitude of the swept-sine for the diode clipper as shown in Figure 2.3, with the amplitude changed over time. `https://bholmesqub.github.io/thesis/chapters/circuit-modelling-identification/`

stages to capture the behaviour of a tube-based effects pedal. Following this work, the model was reformatted to use two nonlinear stages staggered in series with three linear stages, beginning with the linear stage [68]. These choices were made to better resemble the topology of a guitar amplifier – i.e. the selected case studies – with a pre-amplifier and power amplifier surrounded by linear filtering. Listening tests revealed success at modelling clean amplifiers, with results deteriorating when distortion became more present.

2.2.2 Direct component measurement

Given the objective of finding physical component parameters, direct measurement of each component would appear to be the obvious choice. An immediate problem lies in the disassembly of the circuit: there is a risk attached in the form of possible damage to the circuit which should be keenly avoided if the pedal is vintage/rare.

Direct measurement of components does not only require the probing of the circuit board, but for maximum accuracy, desoldering of each component. Measuring a typical two terminal component, e.g. a resistor, is most accurate when a single path to exist between measurement terminals – through the component – but in a circuit there are many possible routes between these two points.

Manufacturers may even go as far as to obfuscate their designs. The Klon pedal used black epoxy across the PCB to prevent the reverse-engineering of the circuit [69]. Only one successful instance of removing the epoxy and reverse engineering the Klon is required to find the schematic, but to measure each component the epoxy must be removed each time - a difficult task. The difficulty and ardour of direct component measurement therefore prompts the search for alternate methods.

2.2.3 Comparative metrics for identification

When creating a model to emulate a circuit there is a need of one or metrics to compare the circuit and its model. These metrics can be broadly used for two purposes: within the identification procedure, and for validation of the final model. The discussed black box models in e.g. [64, 65, 61] only utilise this comparison for validation as the models are derived analytically from measurements (e.g. using the nonlinear impulse response as kernels for a Volterra series model). On the other hand, the discussed grey box approaches in e.g. [68] uses an iterative optimisation algorithm to find the best fit to a circuit, and so must use a metric to determine what is the best fit.

The most immediate metric to compare circuit and model is that which compares

both outputs, i.e. for the current signal index n and model parameter set the time domain output error is

$$\epsilon_{\mathrm{t}}(n, \boldsymbol{\theta}) = y(n) - \hat{y}(\boldsymbol{\theta}, n), \quad n = 0, \ldots, N_{\mathrm{s}} - 1, \tag{2.40}$$

comparing the accurate output y to the modelled output $\hat{y}(\boldsymbol{\theta})$ for the given model parameters $\boldsymbol{\theta}$. This produces an error signal that can be used to compare the signals at a given sample. To find a single-value metric as typically required in optimisation, further processing is needed.

Deriving an objective function ξ from ϵ_{t} can be found by summing the error at each signal index n. A function must be applied to prevent the reduction of the value of ξ from negative values, for example

$$\xi(\boldsymbol{\theta}) = \sum_{n=0}^{N_{\mathrm{s}}-1} (y(n) - \hat{y}(\boldsymbol{\theta}, n))^2. \tag{2.41}$$

By applying a squaring of the difference all values will be positive (assuming a real valued signal). This metric is used in e.g. [13] but the value is scaled by $1/N_{\mathrm{s}}$ so signals of different lengths will produce comparable values of ξ.

An equivalent comparison can be performed in the frequency domain, exactly proportional to the value of ξ for the time domain by inspecting the magnitude of the DFT:

$$\xi_{\mathrm{f}}(\boldsymbol{\theta}) = \sum_{n=0}^{N_{\mathrm{s}}-1} (|Y(n)| - |\hat{Y}(\boldsymbol{\theta}, n)|)^2, \tag{2.42}$$

where ξ_{f} is the frequency domain objective function, and $|Y(n)|$ represents the magnitude of the DFT of the signal y. A frequency domain approach is applied in [13] where the amplitude spectrogram is used, comparing the frequency domain match over time.

As briefly mentioned in Chapter 1.1, perceptual metrics have been used for the evaluation of audio quality [9], however for the use in audio circuit identification only an initial study has been performed [10]. In the future these metrics could provide effective methods of producing perceptually high-fidelity models without the need for lengthy listening tests. For now this approach is not investigated in favour of more rigorously tested metrics.

2.2.4 Context for the proposed identification strategies

Having discussed the various circuit models and identification strategies used in recent literature, further context can be provided for the objectives from Section 1.1. It was stated that physical models were the intended output of the circuit identification. This

choice was motivated to be able to capture the full range of behaviour over different input amplitudes and audio parameter values. Black box models typically utilise several measurements to derive sets of kernels to match behaviour over a range of input amplitudes, and then further require interpolation in between said kernels to provide a continuous range. Using a mapping function as in the grey box model from [13] provides a better fit to a change in input amplitude, though modelling audio parameters is still neglected. For physical models, the capture of this behaviour is encapsulated within a single model and the challenge then lies in the design of an identification procedure which produces component parameter values that define a model with accurate input/output behaviour over the desired range of inputs and audio parameters.

The measurement strategy utilised to identify circuits using black and grey box models circumvents the need to disassemble the DUT. Only amplitude levels that are anticipated from the preceding signal chain are necessary in the measurements, removing likelihood of damage to internal circuits.

The work presented in this thesis aims for an approach that combines the main advantage of black-box and grey-box models – i.e. model derivation directly from input-output measurent data – with the main advantage of physical models – i.e. inherent modelling of audio parameters. This combination raises new challenges. The computational complexity of a physical model is related to the number of components in a circuit, particularly the number of nonlinearities in the components present. For black and grey box models this, complexity is largely independent from the device; the same model can be reused for many effects, for instance the amplifier model used in [68]. Derivation of kernels from the swept-sine technique is a direct computation from the input/output relationship, the limitation of signal length being dictated by the memory of the measurement and computational devices used.

No such direct computation exists as of yet for the component parameters of complex circuits. Therefore the application of optimisation algorithms is required to reduce the error between the model and circuit output, which involves repeatedly simulating the circuit model output to compare to the measurement. Should a long signal be used the optimisation process may take prohibitively long periods of time to complete. This requires the design of an excitation signal that contains all the necessary amplitude and frequency information while maintaining a short duration, the design of which is discussed in Section 5.2.1. A time domain objective function is used to maximise the information represented by the single-value metric, as opposed to using e.g. the amplitude response which neglects the phase behaviour.

To begin identifying a circuit model, both an initial circuit topology and set of

nominal components is required. A further challenge for physical models lies in the modelling of non-ideal effects, e.g. parasitic capacitances etc. Black box models are often insensitive to non-ideal effects, incorporating parasitic behaviour of components into the model without the need for a change in model topology. For physical models, the opposite is true: should non-ideal effects be present, they must be localised to a specific component which then must be adapted.

If the only objective is to recreate the input/output behaviour of the circuit it is possible that other components in the model may compensate for this non-ideal behaviour, but to what degree will likely depend on components and circuit topology. This approach is taken in Chapter 6 where it is investigated how successful a relatively simple physical circuit model can be calibrated to model the input/output behaviour of a circuit without considering non-ideal effects.

On the other hand, if the exact values of the component parameters is the objective, non-ideal behaviour may lead to error in those values preventing successful identification. Estimation of parameter values is the objective in Chapter 7 where prior to circuit measurements, first the measurement equipment is analysed to include the relevant measurement circuitry into the physical circuit model in an attempt to mitigate its effects.

Chapter 3

Root finding algorithms for nonlinear physical circuit models

In the simulation of physical circuit models, the algorithm with the greatest variation in computational expense is that used to solve the transcendental equations caused by nonlinear component models. For the Nodal-DK method and similar paradigms, these algorithms are essential in the simulation of any circuit featuring nonlinear behaviour. Depending on the selected algorithm and nonlinear component model, computational issues may be encountered: the algorithm may not converge to a solution – or not converge within the time constraint imposed by the computer – causing the model to fail. It follows that determining an algorithm that can consistently solve said equations is critical.

Non-convergence at run-time can be prevented by pre-solving the nonlinearity and storing solutions in a Look-Up Table (LUT). With no time constraint, more computationally-expensive methods can be applied, for example homotopy (further discussed in Section 3.2.4) [30].

While this approach removes the possibility of model failure at run-time, memory costs can be prohibitive. For the Nodal DK method, a naïve (i.e. without approximation) LUT must first be of the same dimensionality as the nonlinear equations with an additional dimension required for each variable parameter present in the model.

Through approximations of a circuit model, LUT size can be reduced. For a guitar preamp with 4 nonlinearities and 3 variable parameters, some LUTs could be reduced to 3 dimensional, with the largest simplified LUT being 6 dimensional [70]. Further simplification using interpolation reduced memory requirements but the problem is not removed.

For this reasonf this chapter focusses on algorithms that attempt to solve nonlinear-

ities at run-time. An algorithm is deemed to be 'robust' if it can successfully solve the required nonlinear equations formed by the circuit models within the given constraints of the computer, i.e. time, arithmetic representation. Example circuits models are simulated with a variety of solvers to compare the robustness and computational-efficiency of each algorithm. Existing iterative solvers are discussed in Section 3.2, noting those used already within the field of VA. Given that the nonlinear equations of the Nodal DK method have a common form, two algorithms are presented that utilise this information. These two algorithms are presented in Section 3.3, with their comparison to the existing methods for the simulation of VA circuits investigated in Section 3.4.

3.1 Case studies

The two circuits selected for the comparison of algorithms are the diode clipper and common-emitter amplifier. These circuits use semiconductor elements, the diode and BJT, which both consist of *pn*-junctions modelled with exponential terms. The algorithms are therefore tested on univariate and multivariate cases of similar nonlinearities, providing some level of comparability of results between case studies. Further, as the derivative of an exponential is itself an exponential, equations formed of exponentials are difficult for gradient-based algorithms to solve – and difficult for other solvers due to the rapid rate of change – thus challenging each algorithm in the comparison.

3.1.1 Diode clipper

The diode clipper sub-circuit is commonly found in distortion, overdrive, and 'tube screamer' effects, and has been used as a case study for countless physical circuit modelling publications dating back to [18] (and likely further). The basic function of the diode clipper is to limit the voltage applied across the input terminals to the circuit, causing a soft-clipping distortion effect at the output.

Diode clipper circuits can take numerous different forms; Figure 3.1 shows three different versions. Common between each version is an RC sub-circuit forming a first-order lowpass filter. The resistor also forms an effective nonlinear resistor divider with the diode, where the resistance of the diode changes w.r.t. the voltage across it. As the voltage increases the resistance of the diode drops, thereby reducing the output amplitude, causing a clipping effect.

In (a), the single-sided clipper, the voltage is only limited when it exceeds a positive voltage, allowing negative voltages to pass un-clipped. For the clipper in (b), the

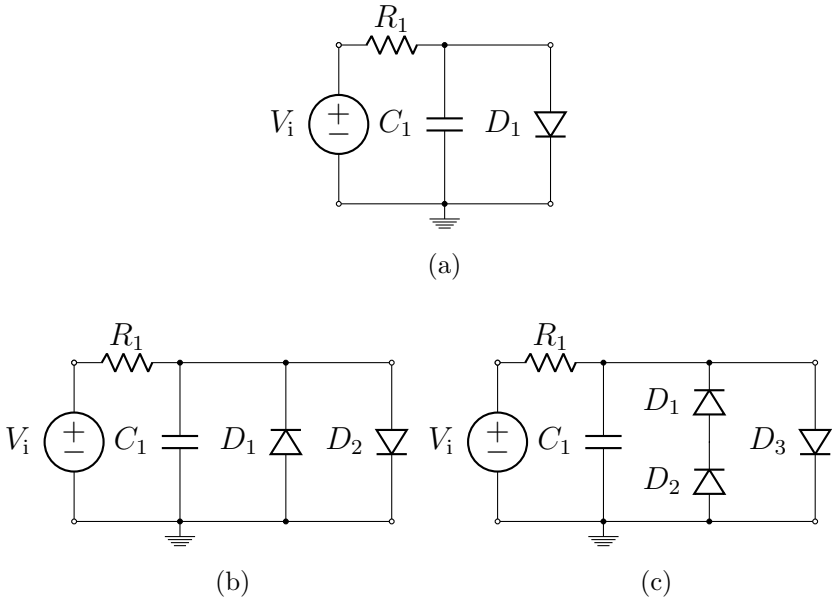


Figure 3.1: Diode clipper circuits: (a) Single sided, (b) Symmetric and (c) Asymmetric.

voltage signal is limited in both positive and negative regions equally. By placing two diodes in series for one direction as in (c), an asymmetry to the clipping is created.

Each of these different versions provides different V-I curves as illustrated in Figure 3.2, which cause different distortion characteristics. More importantly they are modelled using different equations, each are useful for explaining different aspects of methods, and will be used in several examples in this thesis.

Given the Shockley diode model from (2.3), the three versions of the diode clipper nonlinearity can each be modelled by

$$f(v_\mathrm{n}) = I_\mathrm{s}\left(\mathrm{e}^{\frac{v_\mathrm{n}}{NV_\mathrm{T}}} - 1\right) - I_\mathrm{s}\left(\mathrm{e}^{\frac{-v_\mathrm{n}}{\alpha NV_\mathrm{T}}} - 1\right) \tag{3.1}$$

where for the asymmetric diode clipper, $\alpha = 2$, for the symmetric $\alpha = 1$, and for the single sided $\alpha = \infty$.

The asymmetric diode clipper is used in this chapter, with model parameters specified in Table 3.1. Although it is likely for each diode to have different values for I_s, N and V_t, they are set to be the same for simplicity as in this chapter the fit to a real circuit is not the primary concern. Diode parameters are taken from the LTspice model of a 1N4148 diode. The model behaviour was validated against an equivalent simulation from LTspice as illustrated in Figure 3.3.

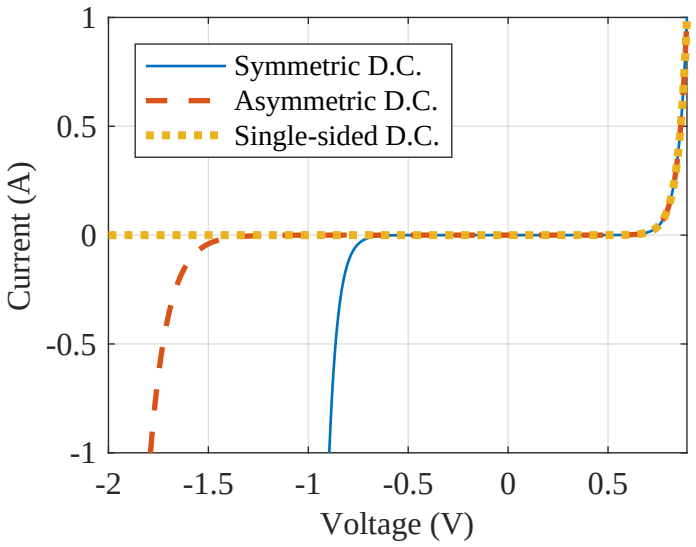


Figure 3.2: Diode clipper V-I curves for symmetric, asymmetric and single sided versions. $I_\mathrm{s} = 2.52\,\mathrm{nA}$, $N = 1.752$, $V_\mathrm{t} = 25.83\,\mathrm{mV}$.

Table 3.1: Parameter values for the asymmetric diode clipper used in the algorithm comparison.

Parameter	Value
R_1	$2.2\,\mathrm{k\Omega}$
C_1	$10\,\mathrm{nF}$
I_s	$2.52\,\mathrm{nA}$
N	1.752
V_t	$25.8\,\mathrm{mV}$

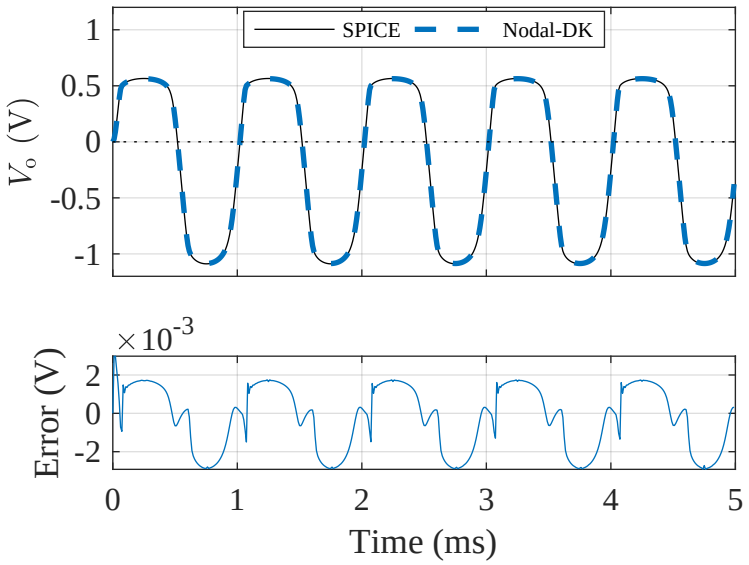


Figure 3.3: Asymmetric diode clipper validation against SPICE, both processing a sine wave with $V_\mathrm{p} = 2\,\mathrm{V}$ and frequency $1\,\mathrm{kHz}$, with the Nodal-DK model simulated with $f_\mathrm{s} = 88.2\,\mathrm{kHz}$. Error is shown in the lower plot to have a peak error of $3\,\mathrm{mV}$.

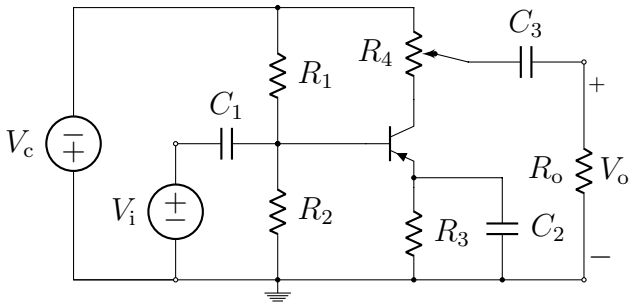


Figure 3.4: Schematic of the common-emitter amplifier.

3.1.2 Common-emitter amplifier

The common-emitter amplifier is commonly used in treble-booster guitar pedals, and also forms a typical building block of analogue audio circuits. Illustrated in Figure 3.4, the common-emitter amplifier is designed to boost the amplitude of the input voltage signal. Analysed from a linear perspective, the circuit forms a combination of a second-order high-pass filter due to the configuration of the capacitors C_1 and C_3, and a shelving filter from C_2 with R_3. If the BJT is driven at a sufficient amplitude, the signal distorts both positive and negative voltages, though with more complex distortion behaviour than that of the diode. The nonlinear equation chosen here for use in the Nodal-DK method consists of the base and collector-current equations of the

Table 3.2: Parameter values for the common-emitter amplifier used in the algorithm comparison.

Parameter	Value
R_1	$470\,\mathrm{k\Omega}$
R_2	$68\,\mathrm{k\Omega}$
R_3	$3.9\,\mathrm{k\Omega}$
R_4	$10\,\mathrm{k\Omega}$
R_o	$1\,\mathrm{M\Omega}$
C_1	$4.7\,\mathrm{nF}$
C_2	$47\,\mathrm{\mu F}$
C_3	$10\,\mathrm{nF}$
I_s	$10\,\mathrm{fA}$
β_f	200
β_r	2
N	1
V_t	$25.8\,\mathrm{mV}$

Ebers-Moll model given in (2.5, 2.6):

$$\mathbf{f}(\mathbf{v}_\mathrm{n}) = \begin{bmatrix} I_\mathrm{b}(V_\mathrm{eb}, V_\mathrm{cb}) \\ I_\mathrm{c}(V_\mathrm{eb}, V_\mathrm{cb}) \end{bmatrix} \quad \text{where} \quad \mathbf{v}_\mathrm{n} = \begin{bmatrix} V_\mathrm{eb} \\ V_\mathrm{ec} \end{bmatrix}, \tag{3.2}$$

recalling that $V_\mathrm{cb} = V_\mathrm{eb} - V_\mathrm{ec}$.

As with the diode clipper this circuit provides a useful case with which to explain different techniques used throughout the thesis. The term 'common-emitter amplifier' refers to the general case, and a specific implementation from the 'Dallas Rangemaster' guitar pedal will be used in the case of vintage circuits. The Dallas Rangemaster used a germanium BJT as opposed to a silicon BJT which is more commonplace today, the difference between the two is discussed further in Chapter 4.

Parameters for the model used in this chapter can be found in Table 3.2. The model behaviour was validated against LTspice, the results illustrated in Figure 3.5.

3.2 Root finding algorithms

As described in Section 2.1.2, the Nodal-DK method forms a set of nonlinear equations of which the root must be found, written as a general function in (2.35). It is assumed

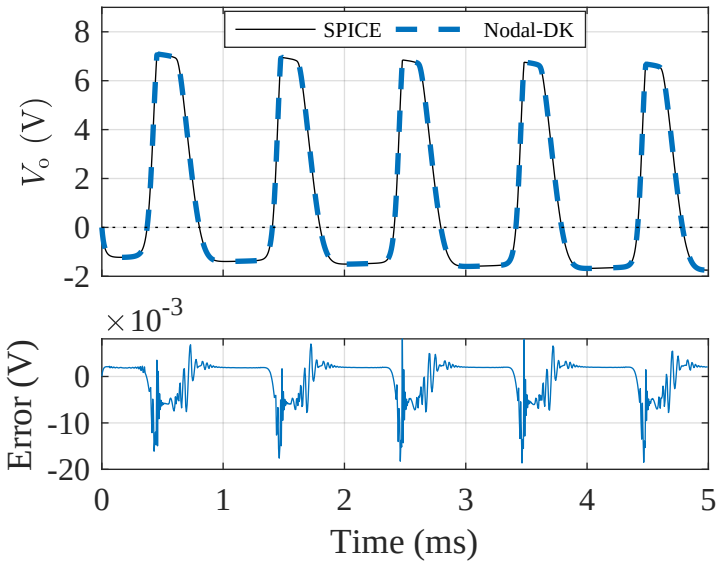


Figure 3.5: Common-emitter amplifier validation against SPICE, both processing a sine wave with $V_\mathrm{p} = 200\,\mathrm{mV}$ and frequency $1\,\mathrm{kHz}$, with the Nodal-DK model simulated with $f_\mathrm{s} = 176.4\,\mathrm{kHz}$. The lower plot shows the error between SPICE and Nodal-DK with a peak of $19\,\mathrm{mV}$.

this equation is transcendental for which a closed-form solution does not exist, and is therefore dependent upon an iterative root finding algorithm to find the solution. The root of the function is defined by

$$\mathbf{g}(\mathbf{v}_\mathrm{n}) = \mathbf{0}. \tag{3.3}$$

With floating point arithmetic, it is unlikely to find a solution at which the function evaluates to exactly zero. The finding of a solution requires convergence specifications, typically defined using the change in values of $\mathbf{v}_\mathrm{n}$ between iterations, i.e.

$$||\mathbf{v}_\mathrm{n}^{i+1} - \mathbf{v}_\mathrm{n}^{i}|| < \mathrm{VTOL} \tag{3.4}$$

where i indicates the current number of iterations and VTOL is the threshold value used to determine convergence. If the magnitude of a solution is not known and a floating point data type is used, a relative value of VTOL can be applied, $\mathrm{RTOL} = \mathbf{v}_\mathrm{n}^{i} \times \mathrm{VTOL}$. This prevents the scenarios where the value of VTOL is smaller than the precision of the data type used to represent the iterate. For the presented case studies, solutions have been found experimentally to fall within a predictable range, removing the necessity of a relative tolerance.

To validate convergence the function can be evaluated at the given solution for proximity to $\mathbf{0}$. Typically, a maximum number of iterations $\bar{i}$ is specified for algorithms

that have not achieved convergence within the given computational requirements. For the demonstration of each method in this chapter, $\mathrm{VTOL} = 1 \times 10^{-12}$ and $\bar{i} = 100$.

Numerous different algorithms can be employed to solve (3.3). In this section there is an overview of notable existing algorithms and also the justification for their inclusion or exclusion from the following analysis. Conditions for inclusion are:

- Ability to extend to multivariate nonlinearities;
- Ability to converge quickly.

The ability to converge quickly is a relative metric for comparing root finding methods. This relates to the convergence rate (e.g. quadratic, linear), but also the computation required for each iteration.

3.2.1 Gradient based solvers

Newton's method

Newton's method forms the base of the root finding method comparison. In recent literature it is frequently chosen for its good convergence properties and ability to extend to multivariate nonlinearities e.g. [59]. A brief derivation of the method is provided as many similar methods are considered for the comparison.

By forming a linear approximation to (3.3) at a given point, a local, closed-form solution can be found. Using the first two terms of the Taylor expansion around $\bar{\mathbf{v}}_\mathrm{n}$, the function is approximated using a linear model

$$\mathbf{g}(\mathbf{v}_\mathrm{n}) \approx \mathbf{g}(\bar{\mathbf{v}}_\mathrm{n}) + \mathbf{J}(\bar{\mathbf{v}}_\mathrm{n})(\mathbf{v}_\mathrm{n} - \bar{\mathbf{v}}_\mathrm{n}), \tag{3.5}$$

where $\mathbf{J}(\bar{\mathbf{v}}_\mathrm{n})$ is the Jacobian (or derivative) matrix of $\mathbf{g}(\mathbf{v}_\mathrm{n})$. From (3.3), a local root can be found by letting $\mathbf{g}(\mathbf{v}_\mathrm{n}) = \mathbf{0}$, leading to the Newton update equation

$$\mathbf{0} = \mathbf{g}(\bar{\mathbf{v}}_\mathrm{n}) + \mathbf{J}(\bar{\mathbf{v}}_\mathrm{n})(\mathbf{v}_\mathrm{n} - \bar{\mathbf{v}}_\mathrm{n}), \tag{3.6}$$

$$\mathbf{v}_\mathrm{n} = \bar{\mathbf{v}}_\mathrm{n} - \mathbf{J}^{-1}(\bar{\mathbf{v}}_\mathrm{n})\mathbf{g}(\bar{\mathbf{v}}_\mathrm{n}). \tag{3.7}$$

This equation is solved iteratively, approximating successive roots of the function until the local root is sufficiently close to the actual root of the function as defined by the convergence criteria. The iterative behaviour becomes clear with a change of variables

$$\mathbf{v}_\mathrm{n}^{i+1} = \mathbf{v}_\mathrm{n}^{i} - \mathbf{J}^{-1}(\mathbf{v}_\mathrm{n}^{i})\mathbf{g}(\mathbf{v}_\mathrm{n}^{i}), \tag{3.8}$$

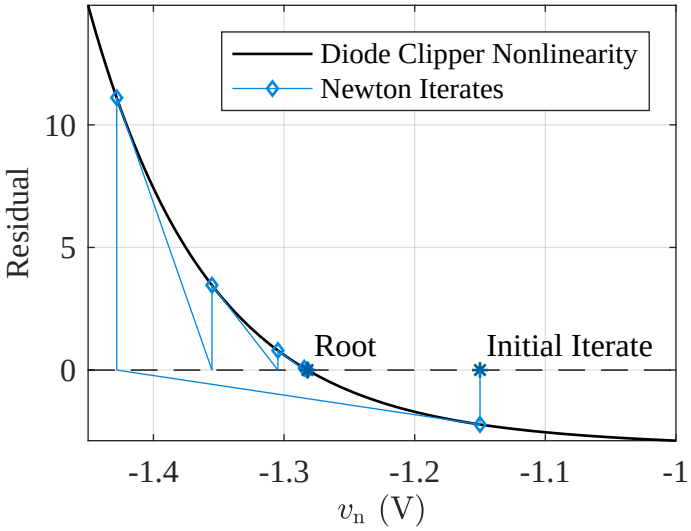


Figure 3.6: Newton's method solving the diode clipper nonlinearity. Overshoot of the method is demonstrated. Each iteration is marked and the gradient followed to the local solution/next iteration. The solution is found within the specified tolerance in 7 iterations.

where the point at which the approximation is formed $\bar{\mathbf{v}}_\mathrm{n}$ now represents the current iterate $\mathbf{v}_\mathrm{n}^i$, and the local solution $\mathbf{v}_\mathrm{n}$ becomes the next iterate $\mathbf{v}_\mathrm{n}^{i+1}$.

If the initial iterate $\mathbf{v}_\mathrm{n}^0$ is sufficiently close to the root $\mathbf{v}_\mathrm{n}^*$, Newton's method can be shown to have q-quadratic convergence, i.e. the residual is roughly squared in the case that the residual is less than 1 (but equivalent behaviour otherwise) [71].

One issue with Newton's method is that it can overshoot the root, leading to the next iteration being further away from the root. This behaviour is illustrated in Figure 3.6 and typically occurs if the curve has an ill-conditioned gradient. As the function is scalar and convex, this overshoot can only occur once [72]. Given the limitations of the floating point representation, even one case of overshoot combined with the exponential functions can cause the following iterate to exceed the maximum representable value, in turn causing the algorithm to fail.

Damped Newton's method

To avoid non-convergence additional conditions can be applied to Newton's method. The Armijo rule, referred to in VA literature as Damped Newton's method [15], describes a line search implemented in the update equation calculation of Newton's

method. Here problematic root-overshoot is defined by the inequality

$$||\mathbf{g}\left(\mathbf{v}_{\mathrm{n}}^{i} - \mathbf{J}^{-1}(\mathbf{v}_{\mathrm{n}}^{i})\mathbf{g}(\mathbf{v}_{\mathrm{n}}^{i})\right)|| > ||\mathbf{g}(\mathbf{v}_{\mathrm{n}}^{i})||, \tag{3.9}$$

stating that the norm of the function at the next iteration is larger than that of the function at the previous iterate. In the event that this inequality is satisfied, a scalar factor is included in the update equation such that

$$\mathbf{v}_{\mathrm{n}}^{i+1} = \mathbf{v}_{\mathrm{n}}^{i} - 2^{-m}\mathbf{J}^{-1}(\mathbf{v}_{\mathrm{n}}^{i})\mathbf{g}(\mathbf{v}_{\mathrm{n}}^{i}), \tag{3.10}$$

where the value of m is the smallest integer that satisfies the inequality [71]

$$||\mathbf{g}\left(\mathbf{v}_{\mathrm{n}}^{i} - 2^{-m}\mathbf{J}^{-1}(\mathbf{v}_{\mathrm{n}}^{i})\mathbf{g}(\mathbf{v}_{\mathrm{n}}^{i})\right)|| \leq ||\mathbf{g}(\mathbf{v}_{\mathrm{n}}^{i})||. \tag{3.11}$$

To implement this in code a `while` loop is used which increments the integer value of m from 1 until (3.11) is satisfied. The behaviour of this algorithm is illustrated in Figure 3.7, where Damped Newton's method is compared to Newton's method on the diode clipper nonlinearity. In the first iteration the step size is reduced by half, dramatically reducing the overshoot. Using the same initial iterate as in Figure 3.6, Damped Newton's method only requires 4 iterations in comparison to the 7 of Newton's method. As Damped Newton's method is compatible with the multivariate case, the improvement in convergence speed justifies inclusion in the algorithm comparison. Due to the additional computation-costs of evaluating the function at each sub-iteration it is possible that the method is less efficient than Newton's method.

Chord method

The Chord method uses an approximation to the Jacobian to reduce the computational cost of each iteration. The Jacobian is only evaluated at the initial iterate, the same values are then used for each successive iteration. In the case of a well-conditioned initial Jacobian and if $\mathbf{v}_{\mathrm{n}}^{0}$ is sufficiently close to $\mathbf{v}_{\mathrm{n}}^{*}$, the Chord method has q-linear convergence [71]. In the case of a poorly chosen initial iterate with an ill-conditioned Jacobian, the method is more prone to increase the residual for each successive iteration without recovery. Both good and poor initial iterate cases are illustrated in Figure 3.8.

Although the Chord method can exhibit poor robustness, by only evaluating the Jacobian once the algorithm is very computationally efficient. For this reason, the algorithm is chosen for the comparison.

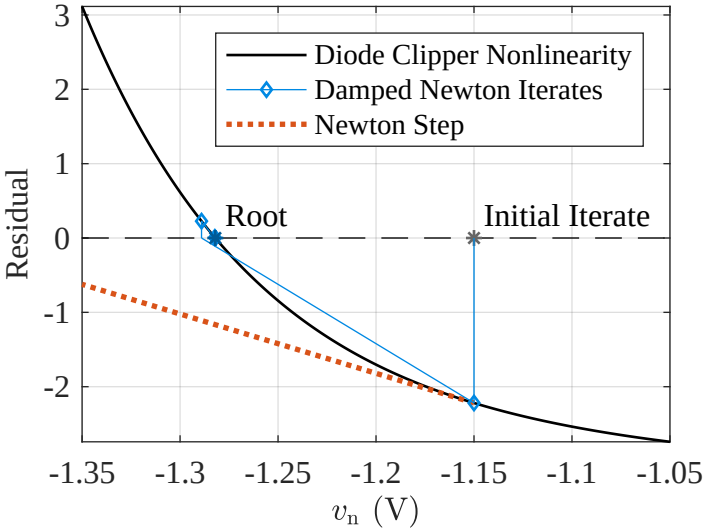


Figure 3.7: Damped Newton's method solving the diode clipper nonlinearity compared to the first step of Newton's method. Damped Newton's method finds the solution within the specified tolerance in 4 iterations.

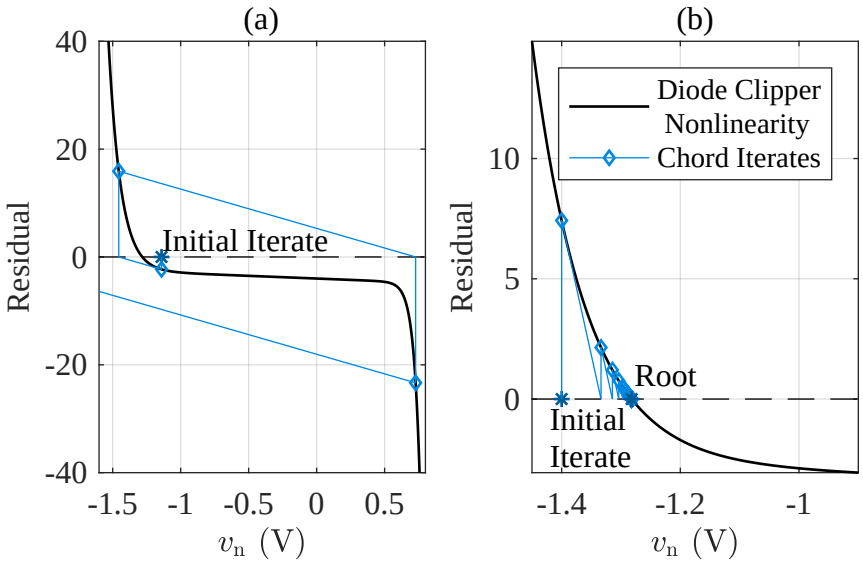


Figure 3.8: Chord method solving the diode clipper nonlinearity. (a) A poor initial iterate, limited to 3 iterations for clarity. (b) A good initial iterate.

Secant method

The Secant method is equivalent to Newton's method except that the Jacobian is calculated using a difference equation. The method has been successfully implemented in VA models for a circuit involving tube nonlinearities [73]. In multivariate cases, the Secant method becomes the Broyden's method, using a difference equation to find the initial Jacobian and an efficient strategy of updating the Jacobian at successive iterates [74]. In an initial investigation, Broyden's method was found to be less robust than the Chord method and so is omitted from the comparison.

Halley's method

Using three terms from the Taylor series (where Newton's method uses two) and solving for the next iterate reveals Halley's method. For the univariate case, this can be expressed by [75]

$$v_{\mathrm{n}}^{i+1} = v_{\mathrm{n}} - \frac{g(v_{\mathrm{n}}^{i})}{J(v_{\mathrm{n}}^{i}) - \frac{1}{2}H(v_{\mathrm{n}}^{i})\frac{g(v_{\mathrm{n}}^{i})}{J(v_{\mathrm{n}}^{i})}} \tag{3.12}$$

where $H(v_{\mathrm{n}}^{i})$ is the second derivative (or Hessian matrix in the case of multivariate nonlinearities) of $g(v_{\mathrm{n}}^{i})$. Though the method extends to multivariate nonlinearities [76], initial tests on the common-emitter amplifier revealed extremely poor convergence and so the algorithm is omitted from the comparison.

3.2.2 Bisection method

A separate family of root-finding algorithms exists which do not utilise the gradient of the function: bracketing methods. Here the bisection method is briefly discussed as a representative example of a bracketing method though due to linear convergence the algorithm is omitted from the comparison.

An interval is defined for which there must be a change of sign in between, i.e. for the interval $[v_{\mathrm{n1}},\ v_{\mathrm{n2}}]$, $\mathrm{sgn}(g(v_{\mathrm{n1}})) = -\mathrm{sgn}(g(v_{\mathrm{n2}}))$. With this condition satisfied, the midpoint of the interval is found. The midpoint then replaces the original boundary with the same sign such that the root is kept within the interval. This routine iterates until the difference between interval boundaries is beneath a set tolerance [77].

Figure 3.9 illustrates an example case of the bisection method finding the root of the diode clipper nonlinearity. For this case, 32 iterations were required to find the solution. Due to the formulation of the algorithm reducing the initial interval size by 2 the number of iterations is decreased by 1. As a result, the convergence can be more

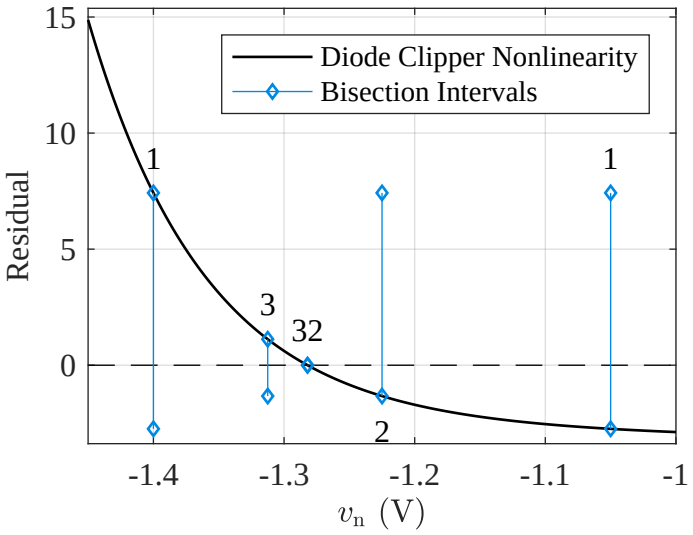


Figure 3.9: Bisection method solving the diode clipper nonlinearity. Intervals are marked with iteration number. Iterations between 4 and 32 have been omitted for clarity.

predictable than that of a gradient-based solver, but due to the high number of function evaluations the algorithm, it is also likely to be computationally intensive.

3.2.3 Lambert W

Though for the methods described in the previous section it is assumed that there is no closed-form solution to the nonlinear equation, one can be found using the Lambert W function for nonlinear equations that can be expressed by the form

$$\mathrm{W}(z)\mathrm{e}^{\mathrm{W}(z)} = z. \tag{3.13}$$

For component models, the Lambert W function can represent the nonlinearity of a diode [78] which has been used in a WDF model [19].

The Lambert W function can also be applied to the Nodal DK method. For a single sided clipper as in Figure 3.1 (a), the Lambert W form can be found through algebraic manipulation of (2.35) such that

$$\mathrm{W}(z) = -\frac{F_\mathrm{n}(f(v_\mathrm{n}) - I_\mathrm{S})}{NV_\mathrm{T}}, \quad z = -\frac{F_\mathrm{n} I_\mathrm{S}}{NV_\mathrm{T}} \mathrm{e}^{\frac{p + F_\mathrm{n} I_\mathrm{S}}{NV_\mathrm{T}}}. \tag{3.14}$$

Solving for $f(v_\mathrm{n})$ then yields

$$f(v_\mathrm{n}) = -\frac{NV_\mathrm{T}}{F_\mathrm{n}}\mathrm{W}(z) - I_\mathrm{S}. \tag{3.15}$$

A symmetric pair of diodes is presented in [19] using the knowledge that only one of the diodes will be forward biased and so will dominate the behaviour of the clipper. This is applied by testing for the polarity of the voltage to detect which diode is dominant, and ignoring the effects of the diode in reverse-bias. For the Nodal-DK method, this is achieved by taking the absolute value of p and multiplying $f(v_\mathrm{n})$ by $\mathrm{sgn}(p)$ to match the correct polarity.

The Lambert W has not been shown to extend to multiple dimensions, and for that reason, it is not compatible with the comparison requirements.

3.2.4 Homotopy

Homotopy describes a process of transitioning between a function with a known solution towards the function of which the solution is desired. Within the field, homotopy has been applied generally to the Nodal-DK nonlinearity to improve the robustness of Newton's method when pre-solving the nonlinearity [30]. This approach used Newton Homotopy as described in [79], where a new function is defined as

$$\mathbf{G}(\mathbf{v}_n^i, \rho) = \mathbf{g}(\mathbf{v}_\mathrm{n}^i) - (1-\rho)\mathbf{g}(\mathbf{v}_\mathrm{n}^0) = \mathbf{0}, \tag{3.16}$$

using ρ to move from the known solution for $\rho = 0$ to the desired solution at $\rho = 1$. Effectively an additional dimension of iteration is required: for each step in the value of ρ, the root-finding algorithm is run to find the solution which is then used for the next value of ρ until the solution for $\rho = 1$ is found. The additional level of iteration leads to a high computational expense which is not suitable for the real-time comparison, but can be useful for other tasks.

Component models are not guaranteed to provide computable models over the full possible range of parameter values, e.g. when $N = 0$ for the Shockley diode. As parameters approach values that cause model-failure, the root finding algorithm faces more challenging scenarios. For this reason, homotopy proves to be an excellent tool in the process of retrieving component parameter values as it is able to solve models over a greater range of parameter values than one of the discussed algorithms alone.

The following chapter focuses on parameter estimation which utilises physically-informed homotopy. SPICE includes an additional, fictional conductance GMIN when solving semiconductor nonlinearities to prevent zero conductances when the junction is reverse-biased (i.e. the model becomes non-computable) [80]. The default value is typically GMIN = 10e-12 which in parallel to a component will have little effect other than to aid convergence. This can be modelled in the Nodal-DK method, for example

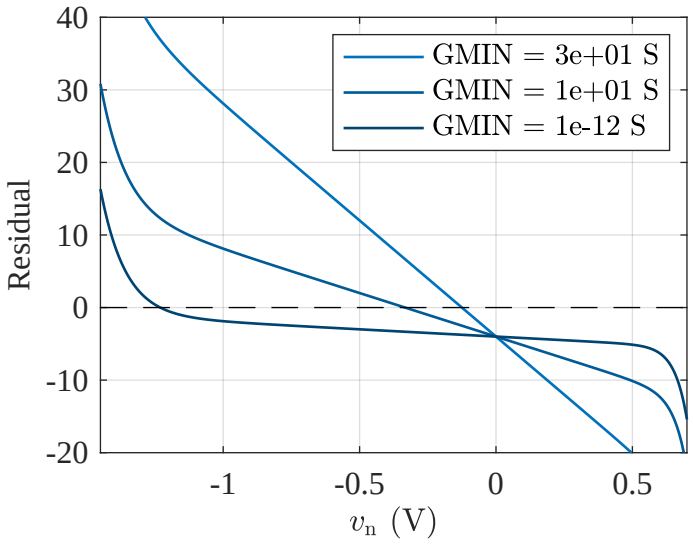


Figure 3.10: Application of homotopy to diode clipper using parallel conductance GMIN.

adding GMIN to the nonlinearity of the single-sided diode clipper

$$0 = p + F_{\mathrm{n}} I_{\mathrm{s}} \left(\mathrm{e}^{\frac{v_n}{N V_{\mathrm{t}}}} - 1 \right) - (1 + \mathrm{GMIN})\, v_n. \tag{3.17}$$

Consider the case in which convergence fails even with the inclusion of GMIN, potentially due to an ill-conditioned gradient at the initial iterate. By increasing, GMIN the gradient becomes less ill-conditioned and the *pn* junction behaves more similarly to a linear resistor.

Illustrated in Figure 3.10 are three choices of a parallel conductance with the diodes in the asymmetric diode clipper. It is clear from the figure that the ill-conditioned gradients in the centre of the illustrated range are transitioned towards well-conditioned gradients by increasing the value of GMIN. By selecting a suitable set of values for GMIN, a transition equivalent to homotopy between a resistor and the *pn* junction can be achieved.

3.3 Deriving system knowledge to improve algorithm performance

The algorithms described so far in this chapter all use features of the function that is being solved: the gradient, the sign etc. Two adaptations are presented here that utilise the form of nonlinearities that arise from the Nodal-DK method to improve robustness and efficiency of (gradient based) root-finding algorithms.

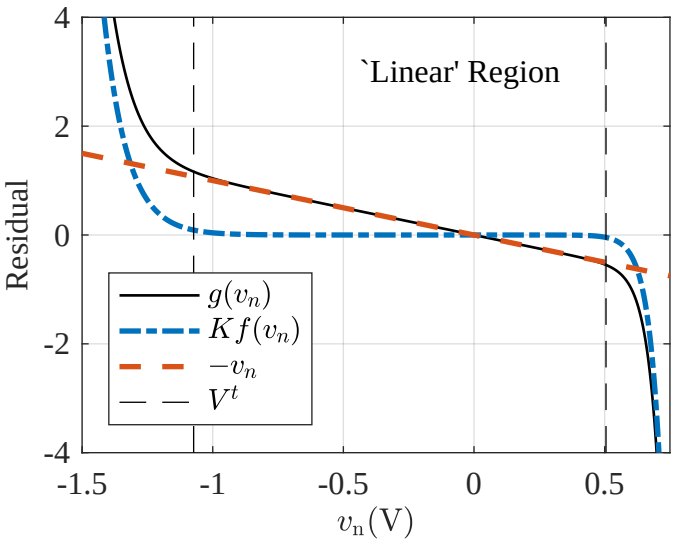


Figure 3.11: Decomposition of the diode clipper nonlinearity into linear and nonlinear regions and their transition voltages V^t.

3.3.1 Derivation from Nodal-DK

From (2.35), the Nodal-DK nonlinearity can be further decomposed into constant ($\mathbf{p}$), linear ($\mathbf{g}_\text{l}$), and nonlinear terms ($\mathbf{g}_\text{n}$),

$$\mathbf{g}_\text{n}(\mathbf{v}_\text{n}) = \mathbf{F}_\text{n}\mathbf{f}(\mathbf{v}_\text{n}), \quad \mathbf{g}_\text{l}(\mathbf{v}_\text{n}) = -\mathbf{v}_\text{n}. \tag{3.18}$$

Figure 3.11 illustrates this decomposition for the asymmetric diode clipper for $\mathbf{p} = \mathbf{0}$. The first method presented in this section utilises the gradient of each term to determine a maximum step-size of a given gradient-based algorithm. The second method uses an approximation to the nonlinearity to provide a strategic initial iterate.

Capped step

From the initial comparison between Newton's method and Damped Newton's method in Figure 3.7, it is clear that Damped Newton's method has desirable convergence properties. This comes at the cost of evaluating the nonlinear function once per sub-iteration of the algorithm. One alternative that requires no in-loop function evaluations would be a hard-limit to the step size of Newton's method. For each element in $\mathbf{v}_\text{n}$, a capped step can be defined by the piecewise function

$$\overline{\Delta v_\text{n}} = \begin{cases} \text{sgn}(\Delta v_\text{n})V^\text{l}, & |\Delta v_\text{n}| > V^\text{l} \\ \Delta v_\text{n}, & |\Delta v_\text{n}| \leq V^\text{l} \end{cases} \tag{3.19}$$

where V^{l} is the limit to the step size Δv_{n}, which becomes $\overline{\Delta v_{\mathrm{n}}}$ after capping. To implement this approach a suitable step size must be found that may improve algorithm robustness.

By treating each term separately (and setting $\mathbf{p} = \mathbf{0}$), the decomposition in (3.18) can be used to establish which term is dominant in different regions of the nonlinearity. The gradient of each term is a possible metric to determine which term is dominant at a given point, and so by finding the points that these gradients are equal provides transitional points, i.e. for the univariate case

$$\frac{dg_{\mathrm{n}}}{dv_{\mathrm{n}}} = \frac{dg_{\mathrm{l}}}{dv_{\mathrm{n}}} \quad \text{when} \quad v_{\mathrm{n}} = V^{\mathrm{t}}, \tag{3.20}$$

where V^{t} is the transitional voltage. It is possible and even likely for there to exist multiple values of V^{t} for a given nonlinearity. The choice from these for the limiting value is left for the user: for the algorithm implemented in this work, one value is found for each dimension, avoiding conditional statements for the selection of a limit within the algorithm loop.

Capping the step size is compatible with each of the described algorithms that operates on one point per iteration (as opposed to the interval used in the Bisection method). As Newton's method forms the baseline to which other adaptations have been applied, it is a clear choice for which algorithm the capped step should also be applied for the algorithm comparison.

Strategic initial iterate

A typical strategy for choosing an initial iterate for the nonlinear equation is to choose the solution from the previous time-step. At low amplitudes/low frequencies this provides a reasonable estimate as the change in root is likely to be small. At high amplitudes/high frequencies the distance between previous and current solutions increases which increases the difficulty to converge to the new solution.

Instead a function could be defined that follows the curve of the nonlinearity, providing an initial iterate in close proximity to the root without the dependence of the solution at the previous time-step. One such function can be determined through algebraic manipulation of (3.18): by excluding the linear term, $\mathbf{f}(\mathbf{v}_{\mathrm{n}})$ can be inverted to find an approximate root $\mathbf{v}_{\mathrm{n}}^{\mathrm{ni}}$,

$$\begin{aligned} \mathbf{0} &= \mathbf{p} + \mathbf{F}_{\mathrm{n}}\mathbf{f}(\mathbf{v}_{\mathrm{n}}^{\mathrm{ni}}) \\ {\mathbf{F}_{\mathrm{n}}}^{-1}(-\mathbf{p}) &= \mathbf{f}(\mathbf{v}_{\mathrm{n}}^{\mathrm{ni}}) \\ \mathbf{f}^{-1}({\mathbf{F}_{\mathrm{n}}}^{-1}(-\mathbf{p})) &= \mathbf{v}_{\mathrm{n}}^{\mathrm{ni}}. \end{aligned} \tag{3.21}$$

This strategy is dependent upon $\mathbf{f}(\mathbf{v}_\mathrm{n})$ being invertible which is true for both the Shockley diode and Ebers-Moll BJT models.

A strategic initial iterate is again compatible with each of the discussed algorithms that operate on a single point of the nonlinearity. To maintain a fair comparison, the new iteration is applied to Newton's method for the algorithm comparison.

3.3.2 Asymmetric diode clipper system knowledge

As the diode clipper nonlinearity is combined from anti-parallel diodes, one direction will always be dominant. Assuming that $\mathrm{e}^{x} \gg \mathrm{e}^{-x}$ is true for the range of the nonlinear equation, an approximate nonlinearity can be derived from (3.1)

$$f(v_\mathrm{n}) \approx \begin{cases} I_\mathrm{s}\left(\mathrm{e}^{\frac{v_\mathrm{n}}{NV_\mathrm{t}}} - 1\right) & v_\mathrm{n} > 0 \\ I_\mathrm{s}\left(\mathrm{e}^{\frac{-v_\mathrm{n}}{2NV_\mathrm{t}}} - 1\right), & v_\mathrm{n} < 0 \end{cases}. \tag{3.22}$$

Capped step

The approximation in (3.22) makes it possible to find two values for V^t that satisfy (3.20),

$$\frac{F_\mathrm{n} I_\mathrm{s}}{NV_\mathrm{t}} \mathrm{e}^{\frac{V^\mathrm{t}}{NV_\mathrm{t}}} = -1, \quad \frac{F_\mathrm{n} I_\mathrm{s}}{2NV_\mathrm{t}} \mathrm{e}^{\frac{-V^\mathrm{t}}{2NV_\mathrm{t}}} = -1, \tag{3.23}$$

both points illustrated in Figure 3.11. From these two values the user may select one to limit the Newton step, though only the magnitude is of interest i.e. $V^\mathrm{l} = |V^\mathrm{t}|$. In this case, the smaller of the two is chosen: for the diode clipper parameters given in Table 3.1, $V^\mathrm{l} = 533.9\,\mathrm{mV}$.

New iterate

Again, using the approximation in (3.22), two solutions to (3.21) can be found for the asymmetric diode clipper:

$$v_\mathrm{n}^\mathrm{ni} = \begin{cases} NV_\mathrm{t}\log\left(1 - \frac{p}{F_\mathrm{n} I_\mathrm{s}}\right), & p \geq 0 \\ -2NV_\mathrm{t}\log\left(1 + \frac{p}{F_\mathrm{n} I_\mathrm{s}}\right), & p < 0 \end{cases}. \tag{3.24}$$

For a given time-step, first p is used to determine which expression to select.

3.3.3 Common-emitter system knowledge

New iterate

To find the inverted form of the Ebers-Moll model it must be decomposed into a matrix-vector product,

$$\begin{bmatrix} I_{\mathrm{b}} \\ I_{\mathrm{c}} \end{bmatrix} = I_{\mathrm{s}} \begin{bmatrix} \frac{1}{\beta_{\mathrm{f}}} & \frac{1}{\beta_{\mathrm{r}}} \\ 1 & -\frac{\beta_{\mathrm{r}}+1}{\beta_{\mathrm{r}}} \end{bmatrix} \begin{bmatrix} \mathrm{e}^{\frac{V_{\mathrm{eb}}}{NV_{\mathrm{t}}}} - 1 \\ \mathrm{e}^{\frac{V_{\mathrm{cb}}}{NV_{\mathrm{t}}}} - 1 \end{bmatrix}. \tag{3.25}$$

The simplified Nodal-DK nonlinearity can then be solved for the exponential terms, such that

$$-\mathbf{Q}^{-1}\mathbf{p} = \begin{bmatrix} \mathrm{e}^{\frac{V_{\mathrm{eb}}^{\mathrm{ni}}}{NV_{\mathrm{t}}}} - 1 \\ \mathrm{e}^{\frac{V_{\mathrm{eb}}^{\mathrm{ni}} - V_{\mathrm{ec}}^{\mathrm{ni}}}{NV_{\mathrm{t}}}} - 1 \end{bmatrix}, \quad \text{where} \quad \mathbf{Q} = \mathbf{F}_{\mathrm{n}} I_{\mathrm{s}} \begin{bmatrix} \frac{1}{\beta_{\mathrm{f}}} & \frac{1}{\beta_{\mathrm{r}}} \\ 1 & -\frac{\beta_{\mathrm{r}}+1}{\beta_{\mathrm{r}}} \end{bmatrix}. \tag{3.26}$$

The voltages used in the Ebers-Moll model have been transformed to ensure they match with those used in the Nodal-DK nonlinearity. Final element-wise operations yield the expressions of the new iterate,

$$\mathbf{v}_{\mathrm{n}}^{\mathrm{ni}} = \begin{bmatrix} V_{\mathrm{eb}}^{\mathrm{ni}} \\ V_{\mathrm{ec}}^{\mathrm{ni}} \end{bmatrix} = \begin{bmatrix} NV_{\mathrm{T}} \log\left(1 - \hat{p}_1\right) \\ NV_{\mathrm{T}} \log\left(\frac{1-\hat{p}_1}{1-\hat{p}_2}\right) \end{bmatrix} \tag{3.27}$$

where $\hat{\mathbf{p}} = \mathbf{Q}^{-1}\mathbf{p}$.

Capped step

As noted previously, the Nodal-DK nonlinearity defined in (3.2) has the Ebers-Moll model using a different set of voltages to those contained in $\mathbf{v}_{\mathrm{n}}$. To simplify the process of determining values for V^{t}, the voltages used for the Nodal DK-method are transformed to match those of the Ebers-Moll model.

$$\hat{\mathbf{v}}_{\mathrm{n}} = \begin{bmatrix} V_{\mathrm{eb}} \\ V_{\mathrm{cb}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \mathbf{v}_{\mathrm{n}}, \tag{3.28}$$

leading to the transformed gradient comparison,

$$\frac{\partial \mathbf{g}_{\mathrm{n}}}{\partial \hat{\mathbf{v}}_{\mathrm{n}}} = \frac{\partial \mathbf{g}_{\mathrm{l}}}{\partial \hat{\mathbf{v}}_{\mathrm{n}}}, \tag{3.29}$$

$$\begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix} = \frac{1}{NV_{\mathrm{t}}} \mathbf{Q} \begin{bmatrix} \mathrm{e}^{\frac{V_{\mathrm{eb}}^{\mathrm{t}}}{NV_{\mathrm{t}}}} & 0 \\ 0 & \mathrm{e}^{\frac{V_{\mathrm{cb}}^{\mathrm{t}}}{NV_{\mathrm{t}}}} \end{bmatrix}. \tag{3.30}$$

Element-wise evaluation of this comparison yields three expressions for V^{t}. Two of these are for V_{eb},

$$V_{\mathrm{eb}}^{\mathrm{t}} = NV_{\mathrm{t}}\log\left(-\frac{NV_{\mathrm{t}}}{Q_{11}}\right) \text{ and } V_{\mathrm{eb}}^{\mathrm{t}} = NV_{\mathrm{t}}\log\left(-\frac{NV_{\mathrm{t}}}{Q_{21}}\right), \tag{3.31}$$

from which as with the diode clipper the smallest value is selected. For V_{cb} one expression is found,

$$V_{\mathrm{cb}}^{\mathrm{t}} = NV_{\mathrm{t}}\log\left(\frac{NV_{\mathrm{t}}}{Q_{22}}\right). \tag{3.32}$$

To transform $V_{\mathrm{cb}}^{\mathrm{t}}$ back to $V_{\mathrm{ec}}^{\mathrm{t}}$, the expression is evaluated at $V_{\mathrm{eb}} = 0$, i.e. $V_{\mathrm{ec}}^{\mathrm{l}} = \mid -V_{\mathrm{cb}}^{\mathrm{t}} \mid$.

3.4 Method comparison

With both the discussed existing algorithms and presented adaptations using information derived from the Nodal-DK method, 5 methods have been selected for comparison:

1. Newton's method;
2. Damped Newton's method;
3. Chord method;
4. Newton's method with a strategic initial iterate;
5. Newton's method with a capped step size.

The three existing algorithms form a range of anticipated robustness from poor (Chord) to excellent (Damped Newton's) with which to compare the algorithms informed from the Nodal DK method. To evaluate the performance of each algorithm first a set of metrics must be determined.

3.4.1 Comparison metrics

While the obvious common metric of performance for each of the selected methods is the number of iterations required to converge, this metric does not account for the difference between the algorithms' computational-cost of each iteration. This motivates the selection of floating point operations (FLOPs) which can be easily evaluated with the help of definitions from the Lightspeed MATLAB toolbox [81]. Results will indicate how the algorithms compare specifically within MATLAB, but should be independent between operating systems. The data type of each variable used is assumed

Table 3.3: Cost of mathematical operations as defined by the Lightspeed Toolbox [81].

Operation	Cost
$+,\ -,\ \times$	1
logical, relational, branch	2
$\mathrm{abs}()$	4
$\mathrm{sgn}()$	5
$\div$	8
$\exp()$	40
$\|\|\mathbf{x}\|\|_2$	$2M + 7$
Solve using LU	$M^3 + \frac{1}{2}M^2 + \frac{29}{2}M - 8$

Table 3.4: Cost in operations of constant values for both diode clipper and common-emitter amplifier models.

Variable	Description	Diode Clipper	Common-emitter
C_{lim}	Step limit calculation cost	32	124
C_{iter}	Strategic iterate calculation cost	37	130
C_{f}	Cost of function evaluation	105	234
C_{j}	Cost of Jacobian evaluation	121	359

to be floating point for simplicity though real algorithms will likely utilise multiple data types. The cost of each operation is displayed in Table 3.3. Branch operations (as they appear in e.g. `if`, `while`) were given the same cost as logical and relational operators.

Using the values and expressions from Table 3.3, the cost of each method is determined based upon the number of dimensions of the nonlinearity M and the number of iterations required to satisfy the convergence condition i. Additionally, the Damped Newton method requires sub-iterations denoted by i_{s}. The costs of calls to the function and Jacobian are represented by C_{f} and C_{j} respectively. C_{lim} and C_{iter} represent the initial cost of calculating the transitional voltages and the approximate initial iterate. These values are found at each time step, assuming that the model is continuously updated due to audio rate parametric control.

The cost of each method is denoted using subscript: C_{N} for Newton's method; C_{D} for Damped Newton's method; C_{C} for the Chord method; C_{CS} for Newton's method

Table 3.5: Model-specific cost in operations for the computation required for one iteration and the initial computation of each method.

	Diode Clipper		Common-emitter	
Method	Initial	Iterate	Initial	Iterate
Newton	234	253	624	646
Damped	234	$261 + 117i_\mathrm{s}$	624	$658 + 252i_\mathrm{s}$
Chord	234	132	624	287
New It.	271	253	754	646
Capped	287	274	790	688

with the capped step applied; and C_NI for Newton's method with the new initial iterate.

$$\begin{aligned} C_\mathrm{N} = {} & M^3 + \frac{1}{2}M^2 + \frac{29}{2}M + C_\mathrm{j} + C_\mathrm{f} - 8 \\ & + i\left(M^3 + \frac{1}{2}M^2 + \frac{35}{2}M + C_\mathrm{j} + C_\mathrm{f} + 8\right) \end{aligned} \tag{3.33}$$

$$\begin{aligned} C_\mathrm{D} = {} & M^3 + \frac{1}{2}M^2 + \frac{29}{2}M + C_\mathrm{j} + C_\mathrm{f} - 8 \\ & + i\left(M^3 + \frac{1}{2}M^2 + \frac{43}{2}M + C_\mathrm{j} + C_\mathrm{f} + 12\right) \\ & + i_\mathrm{s}\left(6M + C_\mathrm{f} + 6\right) \end{aligned} \tag{3.34}$$

$$\begin{aligned} C_\mathrm{C} = {} & M^3 + \frac{1}{2}M^2 + \frac{29}{2}M + C_\mathrm{j} + C_\mathrm{f} - 8 \\ & + i\left(M^3 + \frac{1}{2}M^2 + \frac{35}{2}M + C_\mathrm{f} + 8\right) \end{aligned} \tag{3.35}$$

$$C_\mathrm{CS} = C_\mathrm{N} + 21M + 21iM + C_\mathrm{lim} \tag{3.36}$$

$$C_\mathrm{NI} = C_\mathrm{N} + C_\mathrm{iter} \tag{3.37}$$

Table 3.4 contains the cost of constant values for both the diode clipper and the common-emitter models. Using this information, values were obtained for the cost of an iteration and the initial computation for each algorithm. These are displayed in Table 3.5.

In addition to the conversion from iterations to operations, a moving average filter is applied to the results to simulate the effect of a $2\,\mathrm{ms}$ audio buffer and how it may

process blocks at a time which would smooth the computational load. To illustrate the effects on the iterations/operations, Figure 3.12 shows the unfiltered and filtered iterations of the common-emitter amplifier processing a $30\,\mathrm{ms}$ Hann windowed sine wave at $1\,\mathrm{kHz}$. The signal is simulated at $2\times$ oversampling and solved using Newton's method.

For high-efficiency implementations, VA models will often be written in low-level language such as C or C++ for a target operating system or embedded device. This use case is investigated for the root-finding methods by comparing the average wall clock time of both circuit models written in C. The selected case study circuits using each algorithm are compiled using MATLAB Coder 4.0, an application that converts MATLAB code into C code that can be executed and analysed within MATLAB. Performance may be reduced in comparison to highly optimised C code as the code generated by MATLAB features aspects that enable useful integration with MATLAB, though the performance is much closer to optimal than a MATLAB script.

Time measurements are taken using MATLAB functions `tic` and `toc` which measure wall clock time, and are averaged 100 times to reduce the effects of the other processes being run on the computer. The C code algorithms are executed on a 2017 27-inch iMac with a 4.2 GHz Intel i7 processor and 8 GB of DDR4 RAM clocked at 2400 MHz.

3.4.2 Results

Test simulations were designed to compare the performance of each method against two factors: the amount of oversampling applied, and the peak voltage of the input. Oversampling is compared to test how efficient each method is on computational systems with different processing capabilities.

A 30 period, $1\,\mathrm{kHz}$ sine wave was used to drive the models. The sine wave was modulated by a Hann window so that the amplitude varied across the range of the nonlinearity. For both circuits, the peak voltage of the input was chosen to match what can be expected from a real circuit. As a diode clipper is typically situated after amplification, the highest peak voltage was set at $9\,\mathrm{V}$, which presumes the system uses a dual-rail $\pm 9\,\mathrm{V}$ power supply. To set a voltage range for the common-emitter amplifier it was placed immediately after a guitar as would be the Dallas Rangemaster guitar pedal, so the input reflects a guitar's output. For this reason, a representative maximum peak voltage was set at $300\,\mathrm{mV}$, although it is noted that guitar output voltages can exceed this. The power supply voltage for the common-emitter amplifier model, V_c was set to $9\,\mathrm{V}$.

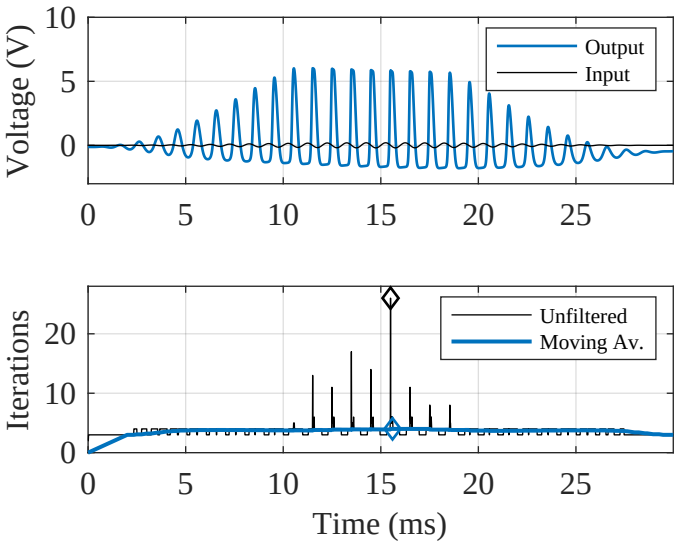


Figure 3.12: Input/Output and iteration count of a $1\ \mathrm{kHz}$, $200\ \mathrm{mV}$ sine wave modulated by a hann window processed by the common-emitter Nodal-DK model using Newton's method, $f_s = 88.2\ \mathrm{kHz}$. Unfiltered and moving average filter results shown, and maximum values marked with $\diamond$.

To ensure a fair comparison, the parameters of the root finding methods were set constant between models and methods, $\mathrm{VTOL} = 10^{-12}$ and $\bar{i} = 100$. Observed inefficiency of Damped Newton's method was corrected by limiting the number of sub-iterations to 3.

Table 3.6 shows results of a set of 16 simulations. Both maximum iteration and operation counts are provided, for which a filtered version and unfiltered version are displayed, along with average simulation time. Figures 3.13 and 3.15 illustrate the performance of the diode clipper and common-emitter amplifier over a range of amplitudes with no oversampling.

The most notable result from these simulations is that both Chord and Newton's methods exhibit non-convergent behaviour in a variety of tests in which the other three methods are convergent. Of these remaining methods, each has cases in which it is the most efficient.

One exclusive feature is the uniform behaviour of Newton's method with a strategic initial iterate. This is clearly observable in Table 3.6 from the consistent behaviour relative to sampling frequency, with the maximum variation of 1 iteration (peak) for the case of the common-emitter amplifier with a peak voltage of 300mV. Figure 3.13 and 3.15 confirm this behaviour relative to input voltage, although with higher variance.

Wall clock time of the simulations with increasing peak voltage are shown in Fig-

ures 3.14 and 3.16 for the SSDC and common-emitter amplifier. For the SSDC, excluding the Chord method, each method is quite insensitive to changes in V_{p}, providing a clear ranking of most efficient to least efficient: Newton's method, Capped step, New Iterate, Damped Newton's method. This contradicts results that focus on FLOPs, in which the New Iterate method is the least efficient. Clearly when simulating in C the cost of calculating the New Iterate is less significant than the per-iteration cost incurred by Damped Newton's method.

For the common-emitter amplifier, the wall clock time trends match closer to those measured in FLOPs, with Damped Newton beginning more efficient than the New Iterate method, but at higher amplitudes the reverse being true. A negative gradient is observed for the New Iterate showing that the larger the amplitude the more efficient it is to utilise the strategic initial iterate. With regards to the results in Table 3.6, as the amount of oversampling is increased, the New Iterate method appears to increase proportionally, whereas Damped Newton's method becomes more efficient at higher sample rates due to benefiting from a closer initial iterate (for example, for the common-emitter amplifier at $300\,\mathrm{mV}$).

Only one case exists at which the C implementation of the Capped Step is less efficient than both Damped Newtons and the New Iterate method, which is for high values of V_{p} for the SSDC. Overall the Capped Step appears to be the most efficient and robust algorithm across each case study and signal parameter (i.e. V_{p} and oversampling). Implementations on different processors may produce different results.

3.5 Conclusion

Two root-finding algorithms were presented using system derived knowledge to improve robustness. Both methods were compared to three existing root-finding algorithms that cover a range of expense/robustness. The results indicate that for cases of moderate peak voltage and higher sampling frequency, Newton's method is sufficiently robust and relatively efficient. However, for more challenging cases (i.e. cases of high peak voltage and/or low sampling frequency), Newton's method was found to be non-convergent. In principle this can be addressed by using Damped Newton's method, although for several tests it proved to be less efficient than both proposed methods.

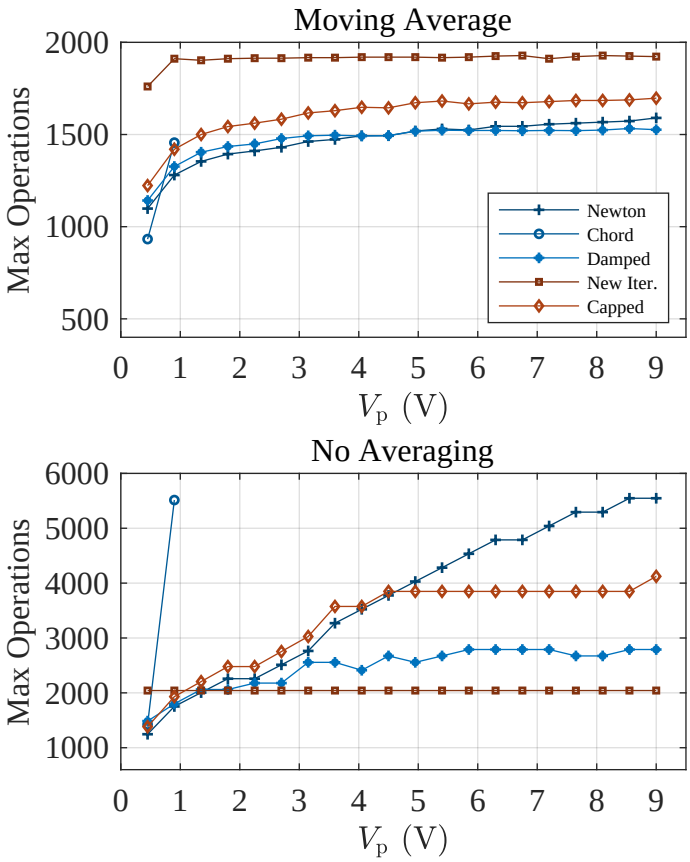


Figure 3.13: Maximum operations against peak input voltage for the asymmetric diode clipper. (Top) Peak averaged operation costs, (Bottom) Peak operation costs.

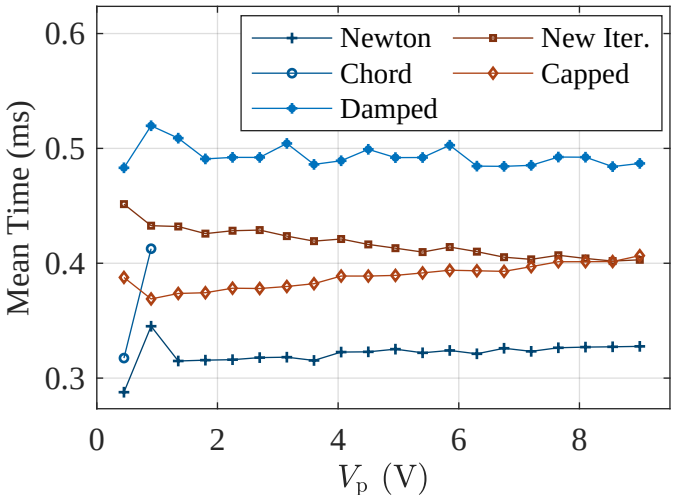


Figure 3.14: Wallclock time of the compared root-finding algorithms against peak input voltage for the asymmetric diode clipper.

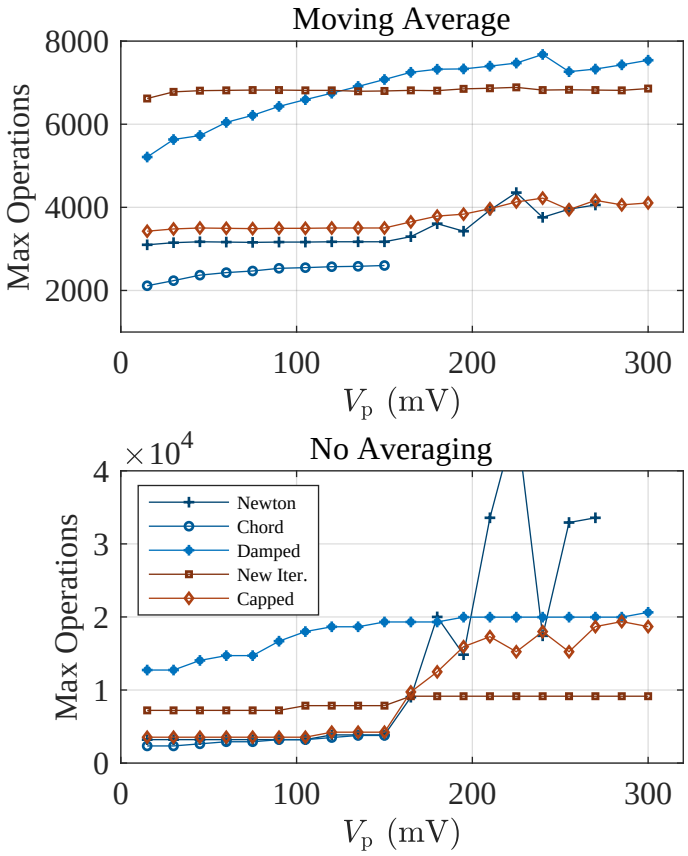


Figure 3.15: Maximum operations against peak input voltage for the common-emitter amplifier. (Top) Peak averaged operation costs, (Bottom) Peak operation costs.

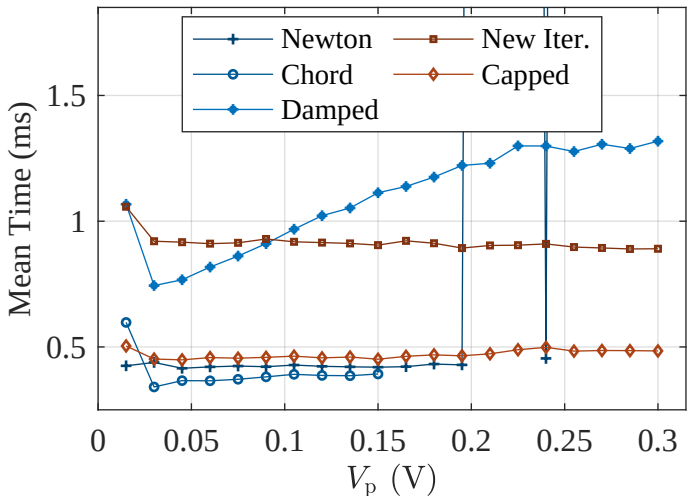


Figure 3.16: Wallclock time of the compared root-finding algorithms against peak input voltage for the asymmetric common-emitter amplifier. Several data points for Newton's method were omitted to improve plot clarity, but are noted here as (V_p, time (ms)): (0.210, 21.004), (0.225, 20.987), (0.255, 24.822), (0.270, 24.631).

3.5.1 Limitations and succeeding work

Utilising predictable behaviour

One avenue of future research is into the removal of the `while` loops from the iterative algorithm. The condition enforced by the `while` loop means that the computer must wait until the end of the loop before knowing what code will be executed next. Modern computer developments enable optimisations to be performed if the program flow is predictable [82].

Newton's method with the new initial iterate exhibited predictable behaviour with a clear limit of iterations over the examined range of operation. Replacing the `while` loop with a `for` loop set to perform the maximum observed number of iterations would enable the computer to know the exact programme flow instead of checking the convergence condition each iteration. It was hypothesised that the optimisations, a compiler could make knowing the exact program flow could outweigh the additional cost of consistently computing the maximum number of iterations. Upon investigation in a C++ programme compiled with maximum speed optimisation (-O3), the `for` loop approach was found to be consistently more expensive than the original `while` loop algorithm.

Circuits with more complex nonlinearities

A key aspect of the proposed iterative methods is that they rely on the availability of an analytic inverse of either the nonlinear term of the equation to be solved for or its first derivative. This criterion is usually satisfied since the components in distortion circuits are often modelled with monotone analytical functions. Even if this holds, to determine the expressions of the capped step and strategic initial iterate several assumptions had to be made about the component behaviour. Should these assumptions break down, it may occur that the methods either do not produce closed form expressions for the transitions and initial iterates, or that the obtained values provide no additional robustness or computational efficiency. Hence a further interesting research direction to explore in future research is to test the methodology on different circuits and more complex component models.

One foray was attempted on this topic, investigating the application of the strategic initial iterate to the Fuzz Face circuit (a circuit description is provided in Chapter 4). Due to the configuration of the two BJTs being connected, the resulting nonlinearity did not work successfully with the proposed strategy of finding a new iteration. The position of the strategic initial iterate causes the Jacobian to become singular within the limitations of the computational system. This scenario could be encountered when

Newton's method using the previous solution as the initial iterate would converge, indicating that the strategic initial iterate does not always provide a robust choice.

Succeeding work

Since the publication of [83] other approaches have been presented, most notably a similar comparison has been performed for WDF [50]. Similar to the Damped Newton's method algorithm, 'Newton's method with backtracking' is coined which implements a line search approach while testing the residual of the next iterate. Additionally a Steepest Descent [77] algorithm is used to find an initial iterate, resulting in significantly reduced iteration counts.

Table 3.6: Results from simulations of both the diode clipper and common-emitter models, $f_s = 44.1\,\mathrm{kHz}$. Average notates a moving average filter has been applied, peak notates no filtering. Entries marked "-" indicate the method was non-convergent.

	$1 \times f_s$					$2 \times f_s$					$4 \times f_s$					$8 \times f_s$				
	Average		Peak		Time	Average		Peak		Time	Average		Peak		Time	Average		Peak		Time
	Its.	Ops.	Its.	Ops.	(ms)	Its.	Ops.	Its.	Ops	(ms)	Its.	Ops.	Its.	Ops.	(ms)	Its.	Ops.	Its.	Ops.	(ms)
Diode Clipper, $V_p = 1\,\mathrm{V}$																				
Newton	3.2	1038	5	1499	0.302	2.8	953	4	1246	0.534	2.6	896	3	993	0.958	2.3	810	3	993	1.750
Damped	3.2	1064	5	1539	1.057	2.8	976	4	1278	1.800	2.6	917	3	1017	3.293	2.3	828	3	1017	7.344
Chord	9.1	1434	48	6570	0.448	5.7	981	15	2214	0.670	4.2	787	8	1290	1.110	3.4	679	6	1026	1.921
New It.	5.5	1670	6	1789	0.429	5.5	1656	6	1789	0.822	5.4	1644	6	1789	1.591	5.4	1631	6	1789	3.097
Capped	3.2	1158	5	1657	0.367	2.8	1066	4	1383	0.666	2.6	1004	3	1109	1.215	2.3	911	3	1109	2.222
Diode Clipper, $V_p = 4.5\,\mathrm{V}$																				
Newton	4.0	1240	13	3523	0.332	3.3	1080	8	2258	0.572	3.0	982	5	1499	0.996	2.7	927	4	1246	1.943
Damped	3.8	1233	7	2295	0.989	3.3	1100	6	1917	1.548	3.0	1005	5	1539	3.551	2.7	949	4	1278	7.483
Chord	-	-	-	-	-	-	-	-	-	-	6.8	1132	99	13302	1.420	4.7	854	17	2478	2.386
New It.	5.5	1667	6	1789	0.455	5.6	1685	6	1789	0.824	5.7	1718	6	1789	1.613	5.8	1742	6	1789	3.351
Capped	4.0	1371	12	3575	0.401	3.3	1203	8	2479	0.685	3.0	1097	5	1657	1.264	2.7	1038	4	1383	2.428
Common-emitter, $V_p = 100\,\mathrm{mV}$																				
Newton	2.9	2518	3	2562	0.421	2.8	2442	3	2562	0.778	2.6	2308	3	2562	1.443	2.3	2091	3	2562	2.681
Damped	5.2	5085	11	12902	0.914	3.9	3619	7	6994	1.445	3.1	2769	5	4670	2.423	2.3	2185	4	3508	4.334
Chord	5.7	2259	8	2920	0.365	4.5	1911	6	2346	0.633	3.8	1703	5	2059	1.153	3.4	1587	4	1772	2.137
New It.	8.4	6176	9	6568	0.890	8.4	6177	9	6568	1.740	8.4	6156	9	6568	3.440	8.0	5922	9	6568	6.657
Capped	2.9	2808	3	2854	0.863	2.8	2726	3	2854	1.698	2.6	2583	3	2854	3.395	2.3	2352	3	2854	6.701
Common-emitter, $V_p = 300\,\mathrm{mV}$																				
Newton	-	-	-	-	-	-	-	-	-	-	2.8	2427	41	27110	1.520	2.5	2241	19	12898	2.858
Damped	6.4	7013	19	23962	1.273	5.0	5122	20	25124	1.932	3.9	3852	23	28610	3.012	3.0	2867	22	27448	5.188
Chord	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
New It.	8.4	6169	12	8506	0.863	8.4	6177	13	9152	1.698	8.4	6151	13	9152	3.395	8.0	5922	13	9152	6.702
Capped	3.8	3434	26	18678	0.478	3.2	3006	21	15238	0.860	2.7	2667	21	15238	1.600	2.5	2510	13	9734	3.056

Chapter 4

Bipolar Junction Transistor modelling for Virtual Analogue

BJTs are one of the most ubiquitous nonlinear components, used in a large number of guitar pedals and amplifiers. It is therefore useful to have available several BJT models with known strengths, from which to choose the most suitable for a given VA circuit model.

The aim of this chapter is to compare a set of nonlinear BJT models in the context of VA circuit models. The baseline of this comparison is the Ebers-Moll model which has already been used extensively in the literature, for example [15, 14, 16]. Since the original Ebers-Moll model was published in 1954 [84], several extensions to the model have been published, notable cases including – in both chronological order and increasing levels of complexity – the Gummel-Poon [85], VBIC [86], and MEXTRAM [87] models. None of the additional effects included in these models has been investigated for use in VA. Here the Ebers-Moll model is extended with additional effects from the Gummel-Poon model resulting in three comparable models.

Note that if a simpler but still nonlinear model is suitable, it is often the case that the model will depend on surrounding circuitry so cannot be examined individually. One case of this is the long-tailed pair which is modelled as a single nonlinearity in models of the Moog ladder filter, e.g. [73].

Nonlinear components such as the BJT are typically the most complex components in a circuit model, and as such, if error is found between a circuit and its model, these components are a likely source of error. A second primary aim arises from this: to find a strategy of extracting model parameters from measurements of a BJT, creating component models as close as possible to real devices to reduce the error in circuit models that use said devices. Isolating the component from the surrounding circuit enables

specific measurements to be performed which are designed for extracting parameters, the objective being to provide more confidence in resultant values.

An interesting challenge is found in germanium BJTs, a semiconductor material that pre-dates the used of silicon and has been used in numerous vintage effects. Due to the lower manufacturing quality when these BJTs were first made and the use of a different semiconductor material, there is potentially more uncharacterised behaviour which would present increased difficulty when fitting models to measurements.

Finally, using the BJT models with extracted parameter values a comparison of VA models is performed: two case study circuits are used to compare the change in model response and computation time. The case studies selected are the Dallas Rangemaster Treble Booster and Arbiter Fuzz Face, both of which originally used germanium BJTs.

4.1 Germanium Bipolar Junction Transistors and circuits

The case studies for this chapter were chosen using two conditions: they must feature one or more BJTs which largely define their behaviour, and the BJTs used must be germanium. This section provides background information on the effects, their design, and which BJT was used.

4.1.1 Dallas Rangemaster Treble Booster

The Dallas Rangemaster Treble Booster is a guitar pedal designed in London circa 1966, and was made famous by guitarists such as Eric Clapton and Jimmy Page [88]. The circuit was designed to sit atop an amplifier, boosting the input signal to cause additional distortion from the amplifier input. The circuit is a common-emitter amplifier as illustrated in Figure 3.4 with the same parameters as in Table 3.2 except for the BJT parameters which will be extracted from measurements. Several BJTs were used in the circuit over the period that the pedal was manufactured, for the analysis in this chapter the OC44 was selected to investigate, its data sheet shown in Appendix A.

One notable difference that is present in the circuit modelled in this chapter as opposed to the circuit in Chapter 3 is that an impedance is placed in series with the input voltage source, illustrated in Figure 4.1. This modification introduces a new signal that drives the circuit $\bar{V}_{\mathrm{i}}$, which after the input impedance R_{i} becomes the original input signal V_{i}.

The introduction of the input resistance is used as a first order approximation of a

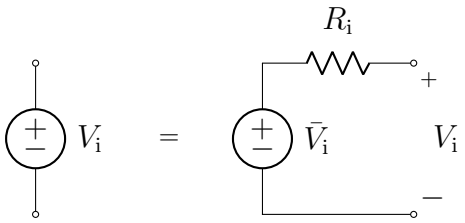


Figure 4.1: Schematic representation of the additional input impedance added to each circuit model to simulate a guitar pickup.

guitar pickup. Guitar pickups use electromagnetic transducers to convert vibrations of a guitar string into an electronic signal. Real pickups exhibit both dynamic and non-linear behaviour. An appropriate equivalent-circuit model of the pickup is presented in [89] which captures the dynamic behaviour of the electronic section of the pickup, separating it from the magnetic section which precedes the equivalent circuit in the model. The objective of the study presented in this chapter is to compare BJT models, and as such introducing the complexity of the pickup in [89] was decided to be a distraction from the primary focus. Cursory comparisons did demonstrate some difference and therefore should readers want to further explore the interaction of a pickup and the presented guitar pedals further and their differences between BJT models, this model would be a good next step.

4.1.2 Arbiter Fuzz Face

The Arbiter Fuzz Face was designed by Ivor Arbiter, again in London circa 1966. It is best known for its use on Jimi Hendrix' 1967 album *Are You Experienced* [90]. Illustrated in Figure 4.2, the Fuzz Face uses two BJTs which form a low component-count Schmitt trigger, typically making the output resemble a square wave. The potentiometer R_f, labelled 'Fuzz', controls feedback from the second BJT stage to the first, filtering the signal with an lowpass created by C_1. The second potentiometer R_v labelled 'volume' forms a potential divider to ground, controlling the amplitude of the output signal.

As with the Rangemaster, several BJTs were used in the pedal. For this analysis, the AC128 has been chosen to remain consistent with previous modelling literature [15], the data sheet can be found in Appendix B. The specific circuit shown in Figure 4.2 was traced from the PCB of a recent iteration of the germanium Fuzz-Face from Jim Dunlop pedals. The parameters for the circuit are shown in Table 4.1.

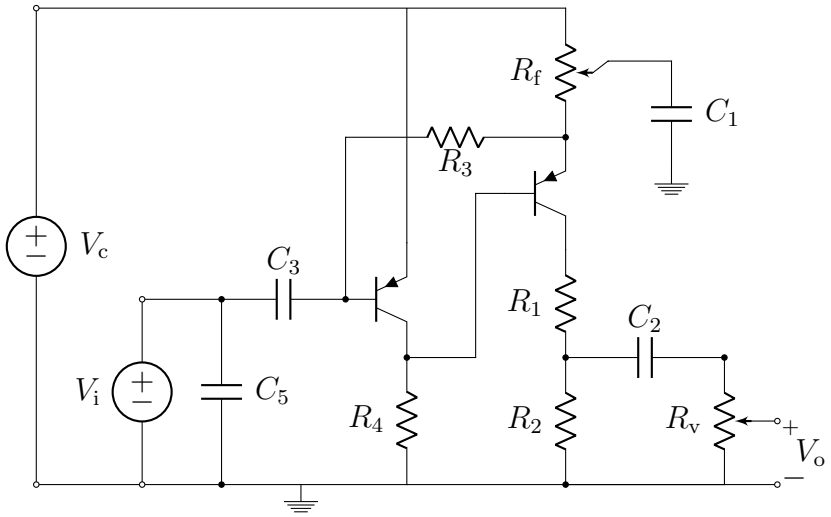


Figure 4.2: Schematic of the Fuzz face circuit.

Table 4.1: Parameter values for the Fuzz Face circuit shown in Figure 4.2.

Parameter	Value
R_1	$11\,\mathrm{k\Omega}$
R_2	$680\,\mathrm{k\Omega}$
R_3	$62\,\mathrm{k\Omega}$
R_4	$36\,\mathrm{k\Omega}$
R_f	$1\,\mathrm{k\Omega}$
R_v	$500\,\mathrm{k\Omega}$
C_1	$22\,\mu\mathrm{F}$
C_2	$0.1\,\mu\mathrm{F}$
C_3	$2.2\,\mu\mathrm{F}$
C_5	$6.8\,\mathrm{nF}$

Table 4.2: BJT regions of operation as defined by junction bias.

	$V_{\mathrm{cb}} < 0$	$V_{\mathrm{cb}} > 0$
$V_{\mathrm{eb}} > 0$	Forward active	Saturation
$V_{\mathrm{eb}} < 0$	Cutoff	Reverse active

4.2 BJT configurations for direct measurement

To discuss the effects of extending the Ebers-Moll model, it is useful to first define the measurement configurations that will be used to characterise each BJT. Four key regions of operation exist for the BJT depending on the junction bias, as shown in Table 4.2 [91]. Of these four states, three are useful in the characterisation strategy presented in this chapter for which three measurement configurations are chosen. A BJT biased in the cutoff regions behaves like an open switch, providing little information about the overall behaviour.

Each of the three selected measurement configurations are illustrated in Figure 4.3:

- Forward Gummel: a measurement of the forward active region, V_{ec} is held positive while V_{eb} is swept and the currents are measured.
- Reverse Gummel: a measurement of the reverse active region similar to the Forward-Gummel but instead V_{ec} is held negative while V_{cb} is swept.
- Common-emitter characteristic: a measurement of the output characteristic, containing both active regions and the transition through the saturation region. The base terminal is driven with a current source and V_{ec} is swept while I_{c} is measured.

The simulated currents of the Ebers-Moll model for each configuration are illustrated in Figure 4.4. From the measured equivalent of this data, the parameter values of the desired models will be extracted. Many additional measurement configurations exist that are designed to extract specific parameters. The three selected configurations largely expose the parameters to be extracted, as will be shown in the following sections.

4.2.1 Silicon vs Germanium BJT comparison

Figure 4.5 shows initial forward Gummel measurements of the OC44, AC128, and 2N3906 BJTs. The 2N3906 is a generic silicon BJT, and though it is not representative

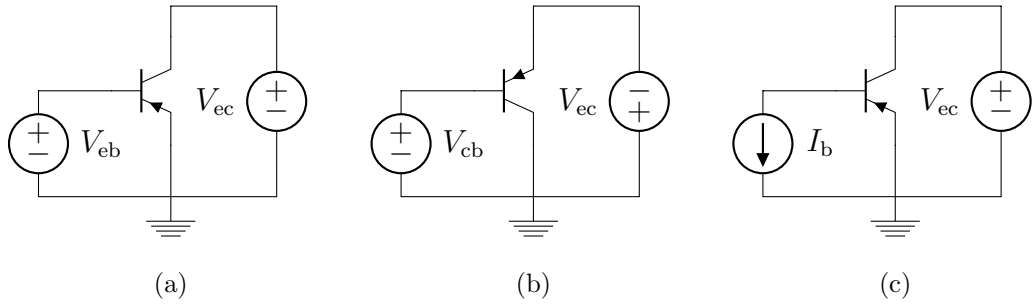


Figure 4.3: Measurement configurations for parameter extraction: (a) forward Gummel, (b) reverse Gummel, (c) common-emitter characteristic.

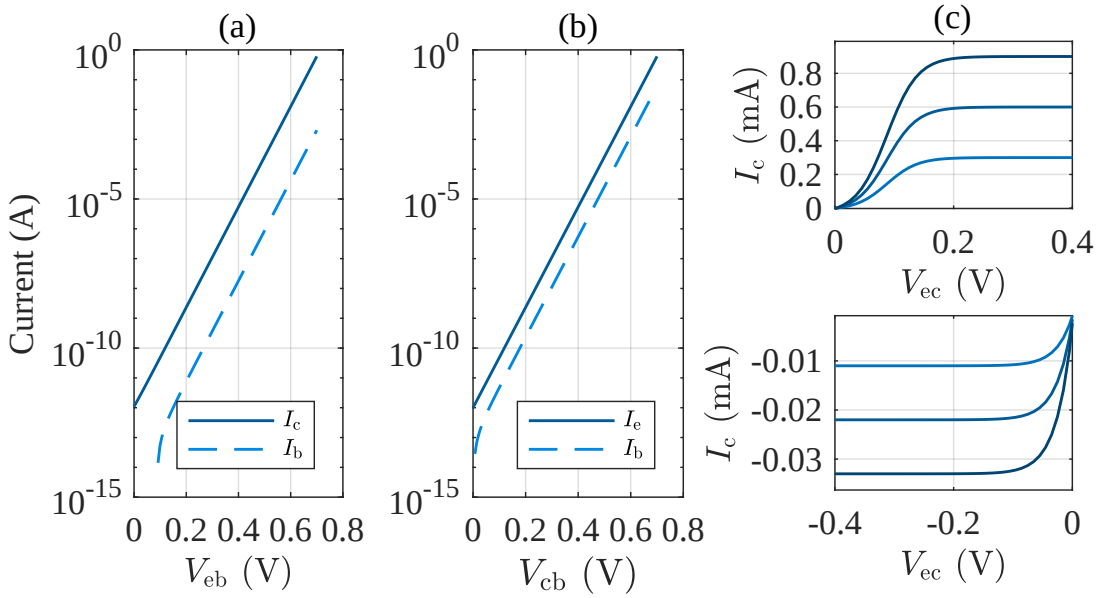


Figure 4.4: Resulting data of measurement configurations modelled with Ebers-Moll: (a) forward Gummel, $V_{ec} = 3\,\mathrm{V}$, (b) reverse Gummel, $V_{ec} = -3\,\mathrm{V}$, (c) common-emitter where $I_b = 1\,\mu\mathrm{A},\ 2\,\mu\mathrm{A},\ 3\,\mu\mathrm{A}$. Model parameters are set as $I_s = 1\,\mathrm{pA}$, $\beta_f = 300$, $\beta_r = 10$, $N_f = N_r = 1$, $V_t = 25.8\,\mathrm{mV}$.

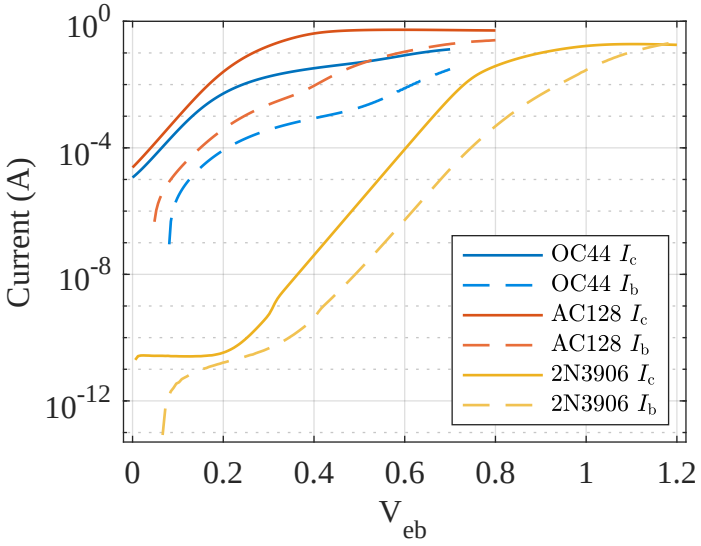


Figure 4.5: Forward Gummel plots of the OC44, AC128 and 2N3906 BJTs measured at $V_{\mathrm{ec}} = 300\,\mathrm{mV}$.

of all silicon BJTs it provides a useful point of comparison.

Recalling the ideal behaviour of a BJT as defined by the Ebers-Moll model and exhibited in Figure 4.4, the 2N3906 BJT exhibits similar behaviour in the middle region of the plot, as can be seen for $300\,\mathrm{mV} < V_{\mathrm{eb}} < 700\,\mathrm{mV}$. These 'ideal' regions are smaller for the OC44 and AC128, closer to $0\,\mathrm{mV} < V_{\mathrm{eb}} < 150\,\mathrm{mV}$, indicating that there may be more demand of additional model complexity for the germanium BJTs.

It must be noted while these measurements allow for comparison between BJTs, they are not suitable for characterisation as they are not sufficiently biased in the forward-active region to isolate the forward-active behaviour.

4.3 Extension of the Ebers-Moll model

To determine whether a more complex BJT will make a significant difference to the output of a VA model, first a set of comparable models must be defined. The Ebers-Moll model forms the baseline with which to compare more complex models. Additional models are extensions of the Ebers-Moll including additional effects based on the Gummel-Poon model [85]. The primary focus of the modelling is on DC effects: the first extended model uses additional terms to fit the high and low current regions, as well as to fit the change in gain with respect to V_{ec}. A third and final model is created by adding junction capacitances, but extraction is not performed on these parameters.

Throughout the derivation of the extended models, repeat references are omitted

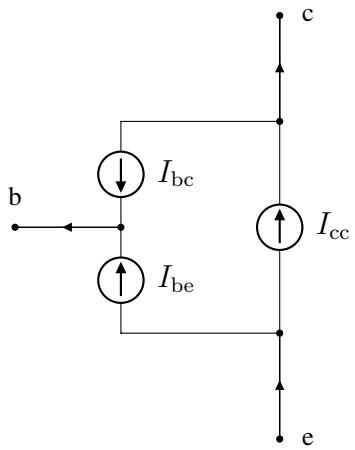


Figure 4.6: Schematic representation of the current source configuration of a PNP BJT.

as the extensions stem from the same reference [85]. Those seeking a more in-depth explanation of the additional effects may refer to [92].

4.3.1 Ebers-Moll revisited

Decomposing the Ebers-Moll model into contributing terms enables an intuitive extension of DC behaviour. The model can be written as a weighted sum of two diodes, beginning with forward and reverse diode terms

$$I_{\mathrm{f}} = I_{\mathrm{s}} \left(\mathrm{e}^{\frac{V_{\mathrm{eb}}}{N_{\mathrm{f}} V_{\mathrm{t}}}} - 1 \right), \quad I_{\mathrm{r}} = I_{\mathrm{s}} \left(\mathrm{e}^{\frac{V_{\mathrm{cb}}}{N_{\mathrm{r}} V_{\mathrm{t}}}} - 1 \right). \tag{4.1}$$

Note that the definition in (2.5, 2.6) only uses one term for the ideality factor N, which has been separated into individual forward and reverse parameters N_{f} and N_{r} for improved versatility in this chapter. The diode terms are combined to form what can be interpreted as VCCSs connected as illustrated in Figure 4.6. Directional notation for the current terms (e.g. bc, base-collector) is chosen to match existing literature that typically refers to NPN BJTs. Three sources are defined as

$$I_{\mathrm{cc}} = I_{\mathrm{f}} - I_{\mathrm{r}}, \quad I_{\mathrm{be}} = \frac{1}{\beta_{\mathrm{f}}} I_{\mathrm{f}}, \quad I_{\mathrm{bc}} = \frac{1}{\beta_{\mathrm{r}}} I_{\mathrm{r}}. \tag{4.2}$$

Finally, the current entering/exiting each terminal is found by examining the contributions of each current source,

$$I_{\mathrm{c}} = I_{\mathrm{cc}} - I_{\mathrm{bc}}, \quad I_{\mathrm{b}} = I_{\mathrm{be}} + I_{\mathrm{bc}}, \quad I_{\mathrm{e}} = -(I_{\mathrm{cc}} + I_{\mathrm{be}}). \tag{4.3}$$

Figure 4.7 shows the forward and reverse Gummel plots of the Ebers-Moll model and how each model parameter is defined assuming ideal behaviour from the BJT.

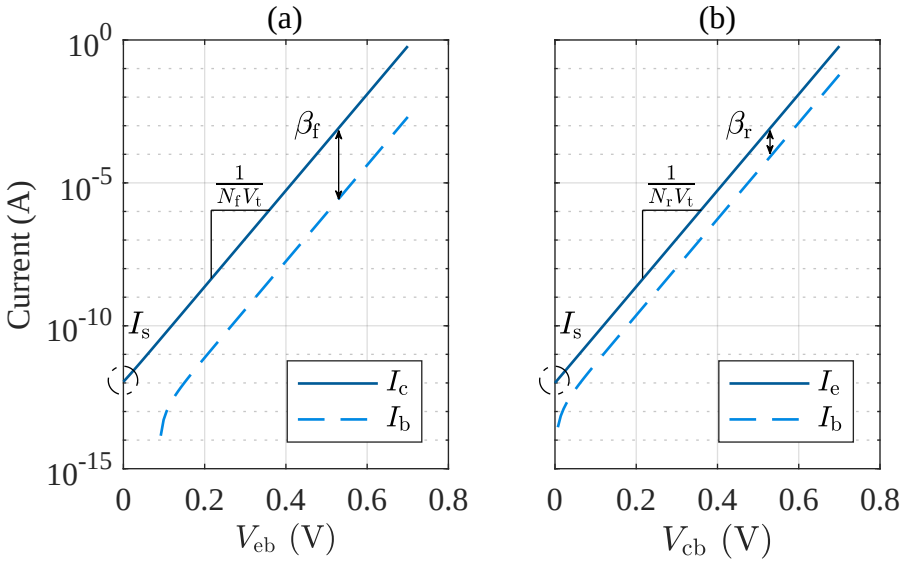


Figure 4.7: (a) Forward and (b) reverse Gummel plots of the Ebers-Moll model. Parameter values are the same as in Figure 4.4.

Saturation current can be extracted using one of two equivalent equations depending on whether the BJT is forward or reverse biased. When forward biased,

$$I_s \approx I_c = -I_s \frac{\beta_r + 1}{\beta_r} \left(e^{\frac{V_{ec}}{N_r V_t}} - 1 \right), \quad V_{ec} \ll 0, \tag{4.4}$$

and when reverse biased,

$$I_s \approx I_e = I_s \frac{\beta_f + 1}{\beta_f} \left(e^{\frac{-V_{ec}}{N_f V_t}} - 1 \right), \quad V_{ec} \gg 0. \tag{4.5}$$

These approximations approach the true value of I_s as β_f and β_r approach infinity, and V_{ec} approaches positive/negative infinity depending on the BJT bias.

Current gains β_f and β_r can be found from constraining the model to their respective active region and finding the relationship between currents, i.e. $\beta_f = I_c/I_b$, $\beta_r = I_e/I_b$.

Ideality factors N_f and N_r define the semi-logarithmic gradient of the Gummel plots, $d\log(I_c)/dV_{eb} = 1/N_f V_t$. To isolate N_f and N_r the temperature of the junction must be measured to find V_t from (2.4).

4.3.2 Early effect

The Early effect describes a dependence of I_c on V_{ec}. This requires the addition of two parameters – the forward and reverse Early voltages V_{af} and V_{ar} – to model the dependency when V_{ec} is both positive and negative. The Early voltages are where

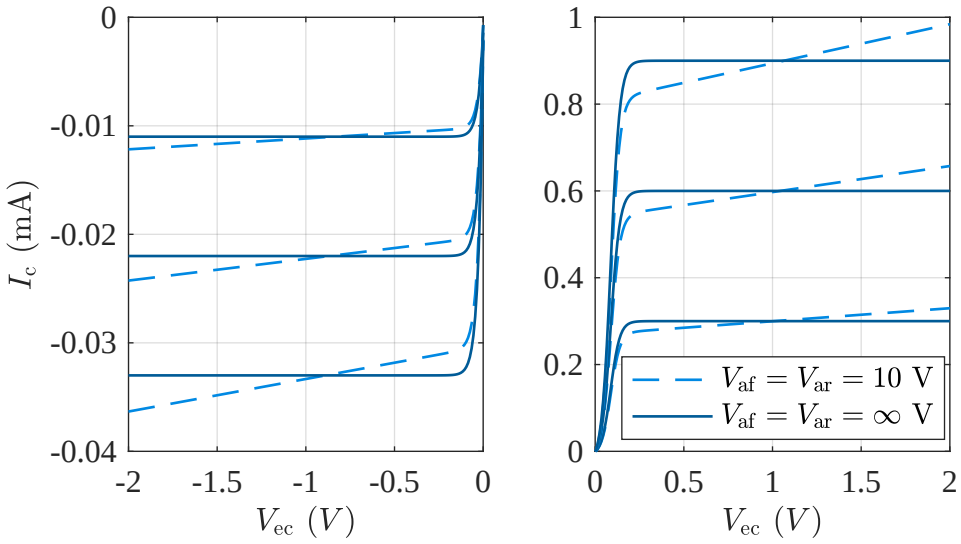


Figure 4.8: Common-emitter characteristic demonstrating the the Early effect. $I_b = 1\,\mu A,\ 2\,\mu A,\ 3\,\mu A$. Ebers-Moll parameter values are the same as in Figure 4.4.

$I_c = 0$ when the curves of the common-emitter characteristic are extrapolated from the active regions. I_{cc} is adapted from (4.2) to include the Early effect with a new term q_1,

$$I_{cc} = q_1^{-1}\left(I_f - I_r\right), \quad q_1^{-1} = 1 - \frac{V_{eb}}{V_{ar}} - \frac{V_{cb}}{V_{af}}, \tag{4.6}$$

which simplifies to the Ebers-Moll model when $q_1 = 1$, i.e. $V_{af} = V_{ar} = \infty$. The Early effect can most clearly be seen on the common-emitter characteristic, illustrated in Figure 4.8. Assuming that the BJT is in forward or reverse active regions, q_1 can be simplified to better indicate how the Early effect depends on V_{ec}. For the forward active region, using $V_{cb} = V_{eb} - V_{ec}$ and assuming $V_{eb} \ll V_{af}$ and $V_{eb} \ll V_{ar}$,

$$q_1^{-1} = 1 - V_{eb}\left(\frac{1}{V_{ar}} + \frac{1}{V_{af}}\right) + \frac{V_{ec}}{V_{af}} \approx 1 + \frac{V_{ec}}{V_{af}}. \tag{4.7}$$

The same approximation can be found for the reverse active region substituting $V_{eb} = V_{cb} + V_{ec}$ and assuming $V_{cb} \ll V_{af}$ and $V_{cb} \ll V_{ar}$,

$$q_1^{-1} = 1 - V_{cb}\left(\frac{1}{V_{ar}} - \frac{1}{V_{af}}\right) - \frac{V_{ec}}{V_{ar}} \approx 1 - \frac{V_{ec}}{V_{ar}}. \tag{4.8}$$

4.3.3 Internal high and low current behaviour

The behaviour of the BJT illustrated on the forward and reverse Gummel plots typically deviates from exponential in two places: high and low currents. This section describes

the extended 'internal' behaviour used to describe the high and low current behaviour, i.e. the additional terms added to the current sources of Figure 4.6. Further high current behaviour is added with the inclusion of external resistances, discussed in the next section.

I_{cc} is extended again from (4.2) to include the term q_2 which models the forward and reverse knee currents I_{kf} and I_{kr},

$$I_{\mathrm{cc}} = \frac{2}{1+\sqrt{1+4q_2}}\left(I_{\mathrm{f}} - I_{\mathrm{r}}\right), \quad q_2 = \frac{I_{\mathrm{f}}}{I_{\mathrm{kf}}} + \frac{I_{\mathrm{r}}}{I_{\mathrm{kr}}}. \tag{4.9}$$

When $q_2 = 0$ (i.e. $I_{\mathrm{kf}} = I_{\mathrm{kr}} = \infty$) the extension simplifies to the Ebers-Moll model. When $I_{\mathrm{kf}} > 0$, I_{c} and β_{f} are reduced at high currents of the forward active region, and equivalently for the reverse active region with I_{e} and β_{r} when $I_{\mathrm{kr}} > 0$.

Low current deviation is modelled by additional exponential terms in I_{be} and I_{bc}, extended from (4.2),

$$I_{\mathrm{be}} = \frac{1}{\beta_{\mathrm{f}}} I_{\mathrm{f}} + I_{\mathrm{se}}\left(\mathrm{e}^{\frac{V_{\mathrm{eb}}}{N_{\mathrm{e}} V_{\mathrm{t}}}} - 1\right), \tag{4.10}$$

$$I_{\mathrm{bc}} = \frac{1}{\beta_{\mathrm{r}}} I_{\mathrm{r}} + I_{\mathrm{sc}}\left(\mathrm{e}^{\frac{V_{\mathrm{cb}}}{N_{\mathrm{c}} V_{\mathrm{t}}}} - 1\right). \tag{4.11}$$

Four new parameters are defined: I_{se}, I_{sc} are the leakage saturation currents, and their corresponding leakage coefficients N_{e} and N_{c}. The additional exponential terms define a second (semi-logarithmic) line segment. Figure 4.9 shows the Gummel plots of the Ebers-Moll model with additional high and low current terms.

4.3.4 Terminal resistances

Terminals and leads of a BJT have an inherent impedance to the flow of current which can be modelled using linear resistors as illustrated in Figure 4.10. The base resistance R_{b} has the largest effect on the high-current region of the forward Gummel plot while the emitter resistance R_{e} effects the reverse Gummel plot, the effect of both can be seen in Figure 4.12. Collector resistance R_{c} controls the transition from the saturation region as illustrated in Figure 4.11. The larger the value of R_{c} the larger the value of V_{ec} must be to forward bias the collector-base junction.

Terminal resistances are not explicitly added to the BJT model but instead are connected between the internal BJT model and external components when implemented in a circuit model. This also applies to the junction capacitances.

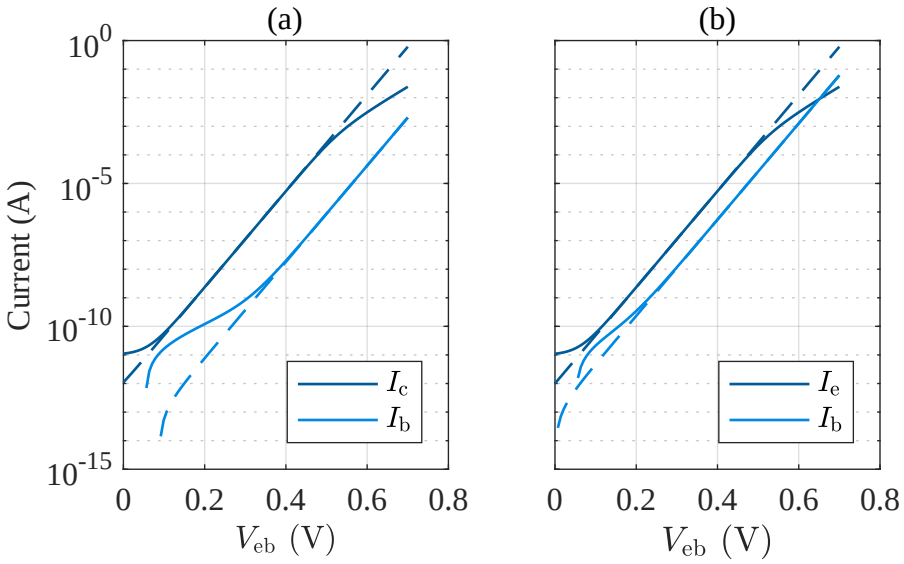


Figure 4.9: (a) Forward and (b) reverse Gummel plots illustrating the difference caused by including the parameters $I_{\mathrm{kf}} = 1\,\mathrm{mA}$ and $I_{\mathrm{kr}} = 1\,\mathrm{mA}$ for the high-current region, and $I_{\mathrm{se}} = 10\,\mathrm{pA}$, $I_{\mathrm{sc}} = 10\,\mathrm{pA}$, $N_{\mathrm{c}} = 3$ and $N_{\mathrm{e}} = 3$ for the low-current region. Dashed lines mark Ebers-Moll model, using the same parameter values as in Figure 4.4, and solid line, the extended model.

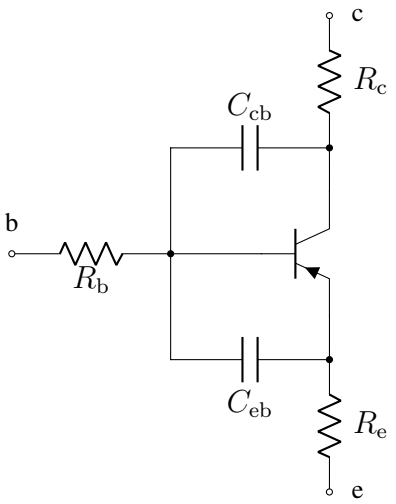


Figure 4.10: Schematic representation of the additional components added to the internal BJT model.

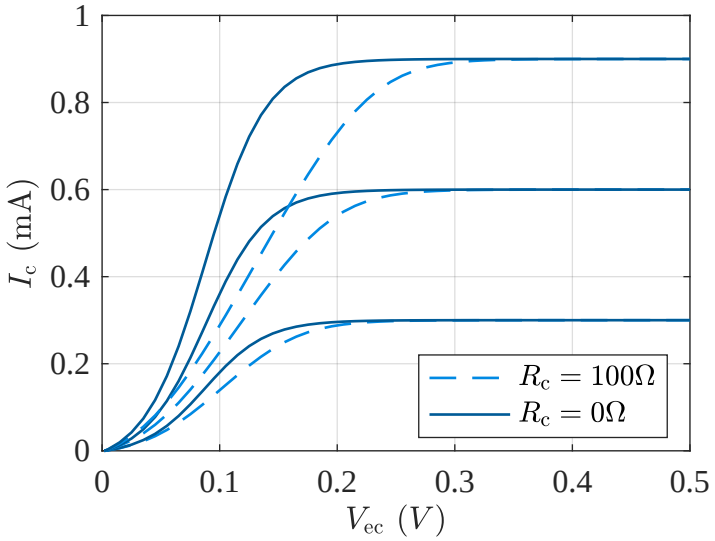


Figure 4.11: Common-emitter characteristic demonstrating the effects of collector terminal resistance $R_c = 100\,\Omega$.

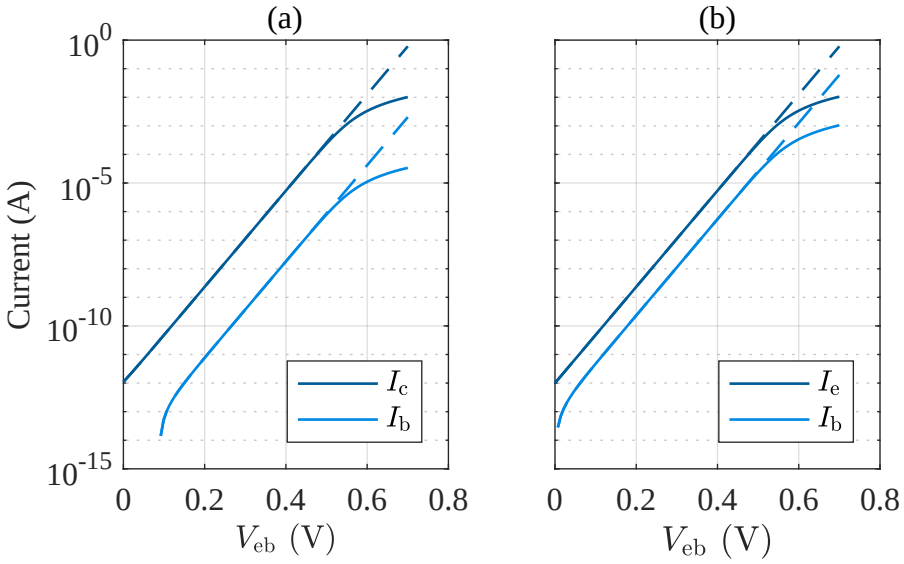


Figure 4.12: (a) Forward and (b) reverse Gummel plots illustrating the difference caused by including the parameters $R_b = 100\,\Omega$ and $R_e = 10\,\Omega$. Dashed lines mark the Ebers-Moll model and solid lines the extended model.

Table 4.3: BJT capacitance values and measurement details given by their datasheets.

BJT	Specification	C_{cb}	C_{eb}
OC44	$V_{\mathrm{ec}} = 6\,\mathrm{V}$, $I_{\mathrm{c}} = 1\,\mathrm{mA}$	$10.5\,\mathrm{pF}$	$410\,\mathrm{pF}$
AC128	$V_{\mathrm{cb}} = -5\,\mathrm{V}$, $I_{\mathrm{e}} = 0\,\mathrm{A}$	$100\,\mathrm{pF}$	-

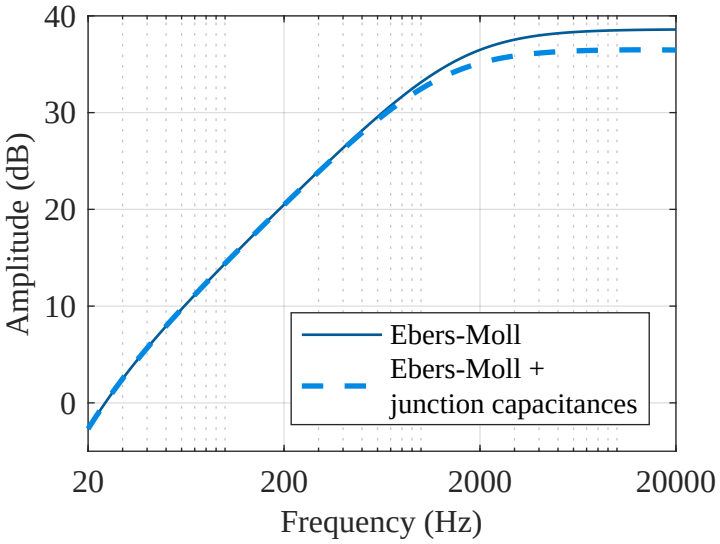


Figure 4.13: Transfer function of the Dallas Rangemaster with and without junction capacitances.

4.3.5 Junction capacitance

In addition to the terminal resistances, Figure 4.10 shows two junction capacitances C_{eb} and C_{cb}. Measuring of dynamic effects requires a significantly different measurement strategy which was not possible with the available equipment. Therefore nominal values are taken from component datasheets given in Table 4.3, provided with specified measurement details.

The effect of the capacitances depends upon the configuration of the BJT in the circuit. For example in the case of the common-emitter amplifier the effective capacitance of C_{cb} is increased due to the inverting gain of the circuit [93]. Figure 4.13 shows the transfer functions of the Dallas Rangemaster circuit with and without junction capacitances. By including these effects the high frequency gain has been reduced by approximately 2 dB.

4.3.6 Compared models

Two extensions to the Ebers-Moll model are defined for the comparison: the DC and AC Gummel-Poon models. While these models are denoted by 'Gummel-Poon' they do not include each term from the original publication [85], but more closely resemble that model than the Ebers-Moll. The internal DC Gummel-Poon model can be expressed by the three VCCS terms:

$$I_{\mathrm{cc}} = \frac{2}{q_1(1+\sqrt{1+4q_2})}\left(I_{\mathrm{f}} - I_{\mathrm{r}}\right), \tag{4.12}$$

$$I_{\mathrm{be}} = \frac{1}{\beta_{\mathrm{f}}} I_{\mathrm{f}} + I_{\mathrm{se}}\left(\mathrm{e}^{\frac{V_{\mathrm{eb}}}{N_{\mathrm{e}}V_{\mathrm{t}}}} - 1\right), \tag{4.13}$$

$$I_{\mathrm{bc}} = \frac{1}{\beta_{\mathrm{r}}} I_{\mathrm{r}} + I_{\mathrm{sc}}\left(\mathrm{e}^{\frac{V_{\mathrm{cb}}}{N_{\mathrm{c}}V_{\mathrm{t}}}} - 1\right). \tag{4.14}$$

The complete DC Gummel-Poon model is defined by (4.12 - 4.14) and the terminal resistances. By further including the junction capacitances the AC Gummel-Poon is defined.

4.4 Parameter extraction of Bipolar Junction Transistors

The proposed BJT parameter extraction strategy is a hybrid direct extraction and optimisation approach based upon two existing strategies [94, 8]. As both were implemented in closed-source commercial software a new approach was created using similar design principles: to initialise optimisation using directly extracted parameter values, and to utilise targeted optimisation stages focussing on reduced groups of parameters.

4.4.1 Measurement details

Measurements were taken using a Keithley 2602B Source Measure Unit (SMU). SMUs are measurement instruments designed for static, DC measurements of devices and are capable of sourcing a voltage or current, whilst simultaneously measuring the remaining quantity, e.g. driving a current and measuring the voltage.

In total, 7 BJTs were measured: 3 OC44 and 4 AC128. Measurements were pulsed to minimise the amount of time the BJT was driven and therefore reduce the change in temperature across measurements. Limits to the measurements – specified in Table

Table 4.4: Ranges of the inputs to each measurement circuit. Specific values of I_b are provided on each measurement plot.

Meaurement	Input	OC44	AC128
Forward	V_eb	0 - 0.7 V	0 - 0.8 V
Gummel	V_ec	2 V	2 V
Reverse	V_cb	0 - 0.8 V	0 - 0.8 V
Gummel	V_ec	–2 V	–2 V
Common	I_b	3 - 50 µA	26 - 1000 µA
Emitter	V_ec	–5 - 5 V	–5 - 5 V

4.4 – were found through setting current limits as noted by the data sheets and finding a maximal voltage that would prevent the SMU from hitting said limit for each BJT.

4.4.2 Direct extraction

Direct extraction of model parameter values is used to find a set with which to initialise the optimisation. Not every parameter requires direct extraction: for this method, only I_s, N_f, N_r, β_f, β_r, I_kf, and I_kr were extracted. A quick manual tuning of the remaining parameters provides initial values that offer a reasonable fit to the three measured characteristics. Extraction procedures were developed from [94].

First V_t must be calculated using (2.4). The temperature T_K represents the temperature of the *pn* junctions which approaches ambient when the BJT is not driven. A measurement of ambient room temperature can therefore be used for the calculation of V_t so long as the junction is not driven for extended periods and is allowed a period of time to return to ambient temperature. Room temperature was measured at $T_\mathrm{K} = 296.45\,\mathrm{K}$ meaning $V_\mathrm{t} = 25.5\,\mathrm{mV}$.

With a value for V_t, values for N_f and N_r can be found. As previously stated, $1/NV_\mathrm{t}$ is the semi-logarithmic gradient of the Gummel plots. As the gradient of the currents in the Gummel plots change with respect to voltage, a suitable measurement point must be selected. The point must be in the ideal region as the extracted parameters belong to the Ebers-Moll model. Using the knowledge from the extended Ebers-Moll model behaviour, we know that the gradient decreases with high-current effects and low-current effects. It then follows that the gradient of the ideal region is the maximum of the curve. Figure 4.14 shows an example gradient curve of I_c in the forward active region, with the extracted gradient value marked with a dashed line.

Table 4.5: Voltage ranges over which each optimisation for both models were performed. Gummel plots were used in both the penultimate and ultimate stages for the Gummel-Poon model, and are labelled 1 and 2 to differentiate.

Model	Measurement	Input	Lower limit	Upper limit
Ebers-Moll	Gummel plots	$V_{\mathrm{eb}}, V_{\mathrm{cb}}$	$10\,\mathrm{mV}$	$200\,\mathrm{mV}$
Gummel-Poon	Current gain	$V_{\mathrm{eb}}, V_{\mathrm{cb}}$	$110\,\mathrm{mV}$	$700\,\mathrm{mV}$
	Gummel plots 1	$V_{\mathrm{eb}}, V_{\mathrm{cb}}$	$100\,\mathrm{mV}$	$700\,\mathrm{mV}$
	Gummel plots 2	$V_{\mathrm{eb}}, V_{\mathrm{cb}}$	$50\,\mathrm{mV}$	$600\,\mathrm{mV}$
	Common-emitter	V_{ec}	$-5\,\mathrm{V}$	$5\,\mathrm{V}$

At this same voltage a value for I_{s} can be extracted. As I_{c} and I_{e} deviate from the ideal exponentials in real devices, instead of following both curves to their intercept, an exponential approximation is made in the ideal region, e.g. for the forward Gummel plot,

$$I_{\mathrm{c}} = I_{\mathrm{s}}\,\mathrm{e}^{\frac{V_{\mathrm{eb}}}{N_{\mathrm{f}}V_{\mathrm{t}}}}, \quad I_{\mathrm{s}} = I_{\mathrm{c}}\,\mathrm{e}^{-\frac{V_{\mathrm{eb}}}{N_{\mathrm{f}}V_{\mathrm{t}}}}. \tag{4.15}$$

Gain parameters use a different point in the measurement to extract a value from: the maximum of the current ratios. The gain extraction point is illustrated for the forward Gummel plot in Figure 4.15. This selection again follows from the logic that the maximal gain of the BJT will be where low- and high-current effects are not reducing its value. Though different extraction points are used for different Ebers-Moll parameters, the extracted values are only used to seed the optimisation process: final parameter values will have been optimised to best fit the measured behaviour of the BJT.

Finally, values for I_{kf} and I_{kr} are found by inspecting the current ratios and finding the values of I_{c} and I_{e} at which the ratios are half of extracted current gains, i.e.

$$I_{\mathrm{kf}} = I_{\mathrm{c}} \quad \text{where} \quad \frac{I_{\mathrm{c}}}{I_{\mathrm{b}}} = \frac{1}{2}\beta_{\mathrm{f}}, \tag{4.16}$$

$$I_{\mathrm{kr}} = I_{\mathrm{e}} \quad \text{where} \quad \frac{I_{\mathrm{e}}}{I_{\mathrm{b}}} = \frac{1}{2}\beta_{\mathrm{r}}. \tag{4.17}$$

Should the current ratio not fall to half of the extracted value of β, the measurement can be extrapolated to find this point.

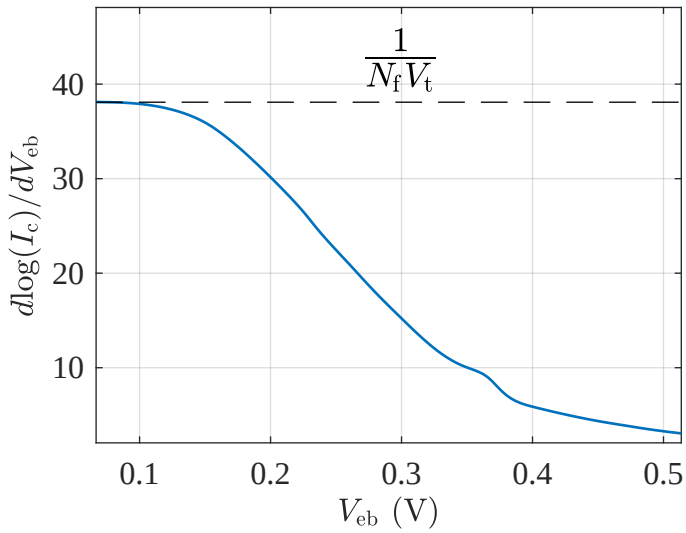


Figure 4.14: A example plot of $\frac{d\log(I_c)}{dV_{eb}}$ illustrating at which point extraction is performed for N_f.

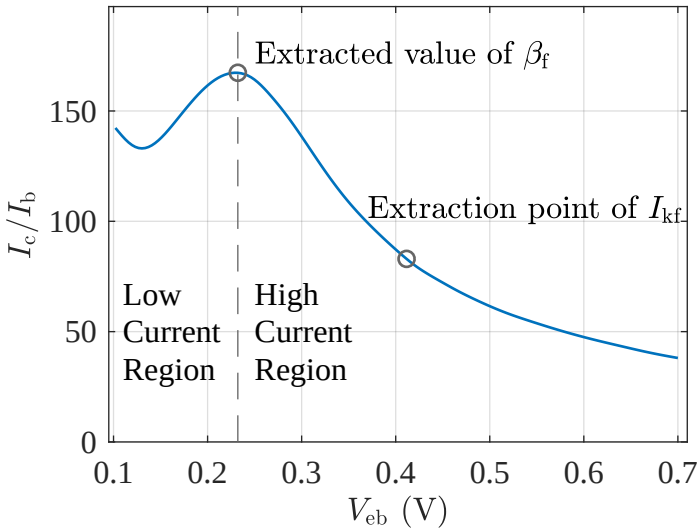


Figure 4.15: An example plot of the ratio I_c/I_b indicating how the gain of the BJT changes with respect to V_{eb} and at which point β_f and I_{kf} are extracted.

4.4.3 Extraction using optimisation

To improve the fit of the BJT models with extracted and manually tuned parameters, a set of targeted optimisations were performed. An error function is defined,

$$\epsilon(\boldsymbol{\theta}_\mathrm{b}, y) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{y(n) - \hat{y}(\boldsymbol{\theta}_\mathrm{b}, n)}{y(n)}\right)^2, \tag{4.18}$$

where $\boldsymbol{\theta}_\mathrm{b}$ are the parameters for a specified BJT model and $y(n)$ is the relevant data at time index n. Observed current values of germanium BJTs can vary over 6 orders of magnitude; normalising each error value by $y(n)$ weights the function to ensure that error when $\hat{y}$ is small is not under-represented.

Four error functions exist for different data: gain/current ratios ϵ_cr, forward Gummel ϵ_fg, reverse Gummel ϵ_rg, and common-emitter characteristic ϵ_ce. Objective functions are created from the combinations of data:

- Gain/current ratios: $\xi_\mathrm{cr}(\boldsymbol{\theta}_\mathrm{b}) = \epsilon_\mathrm{cr}(\boldsymbol{\theta}_\mathrm{b}, y)$;
- Gummel plots: $\xi_\mathrm{g}(\boldsymbol{\theta}_\mathrm{b}) = \epsilon_\mathrm{fg}(\boldsymbol{\theta}_\mathrm{b}, y) + \epsilon_\mathrm{rg}(\boldsymbol{\theta}_\mathrm{b}, y)$;
- Gummel plots and common-emitter characteristic: $\xi_\mathrm{g,ce}(\boldsymbol{\theta}_\mathrm{b}) = \epsilon_\mathrm{ce}(\boldsymbol{\theta}_\mathrm{b}, y) + 0.01 \times \xi_\mathrm{g}(\boldsymbol{\theta}_\mathrm{b})$.

A heuristic weighting is applied in the combined Gummel/common-emitter characteristic objective function to account for the increased number of measurements that the common-emitter characteristic uses.

Two optimisation algorithms were used from MATLAB's optimisation toolbox, `fminsearch` which uses the Nelder-Mead simplex method [95], and `fmincon`, which uses the interior-point method [96]. The Nelder-Mead simplex method is useful in this scenario due to its ability to handle discontinuous surfaces. This enabled the use of objective functions that would return an infinite value if the parameters supplied were negative, preventing non-physical parameter sets. Experimentally, it was found that this combination provided better convergence properties than using the interior-point method with a similar boundary. In optimisation stages where more complex boundaries were required to ensure a suitable starting point was found for the following stage, the interior-point method was chosen. The final stage of characterising each model was performed with the Nelder-Mead simplex method.

As mentioned in Section 3.2.4, a physically informed homotopy solver was implemented for solving the common-emitter characteristic when terminal resistances were

Table 4.6: List of all parameters, constraints used in the intermediate optimisation stages, and initial values used for parameters that were not found through direct extraction.

	Parameter	Optim. Constraints		Init. Values
		Lower Lim.	Upper Lim.	
I_s	Saturation current	-	-	-
β_f	Forward current gain	50	250	-
β_r	Reverse current gain	3	20	-
N_f	Forward ideality factor	-	-	-
N_r	Reverse ideality factor	-	-	-
(V_t)	Thermal voltage	-	-	-
V_af	Forward Early voltage	-	-	-
V_ar	Reverse Early voltage	-	-	-
I_kf	Forward knee current (gain roll-off)	$10\,\mu\mathrm{A}$	$500\,\mathrm{mA}$	-
I_kr	Reverse knee current (gain roll-off)	$10\,\mu\mathrm{A}$	$500\,\mathrm{mA}$	-
I_se	BE junction leakage current	$0.1\,\mathrm{fA}$	$1\,\mathrm{mA}$	$I_\mathrm{s}/2$
I_sc	BC junction leakage current	$0.1\,\mathrm{fA}$	$1\,\mathrm{mA}$	$I_\mathrm{s}/2$
N_e	BE junction leakage emission coefficient	0.5	4	2.3
N_c	BC junction leakage emission coefficient	0.5	4	2.4
R_b	Base resistance	$1\,\Omega$	$250\,\Omega$	$25\,\Omega$
R_e	Emitter resistance	$0.1\,\mathrm{n}\Omega$	$2\,\Omega$	$10\,\mathrm{m}\Omega$
R_c	Collector resistance	-	-	$10\,\mathrm{m}\Omega$
C_eb	Emitter-base capacitance	-	-	-
C_cb	Collector-base capacitance	-	-	-

used. Values of GMIN ranged from $1\,\mathrm{kS}$ to $10\,\mathrm{zS}$ with 100 points spaced logarithmically. The value of GMIN was decreased according to the number of iterations required for Damped Newton's method to converge. For the number of executed iterations i to reach a solution, and the maximum number $\bar{i}$,

- $i < \bar{i}/10$ then GMIN moved 10 values,
- $i < \bar{i}/5$ then GMIN moved 5 values,
- $i < \bar{i}/3$ then GMIN moved 3 values,
- $i < \bar{i}/2$ then GMIN moved 2 values,
- otherwise, GMIN incremented by one value.

Ebers-Moll

Following the direct extraction, one stage of optimisation is applied to find the final values of the Ebers-Moll model. The optimisation uses both Gummel plots with a low voltage range (see Table 4.5) to match the first 'ideal' region in which the gradient of the collector current is approximately constant. Final parameter values can be found in Table 4.7.

DC Gummel-Poon

The optimisation procedure for the DC Gummel-Poon model is illustrated in Figure 4.16. After the direct extraction stage, three stages of optimisation are used. The intermediate optimisation stages use the interior-point method with constraints as shown in Table 4.6.

The first optimisation stage works on the current gain of the BJTs, significantly reducing the effects of I_s, V_af, V_ar, R_c, N_f, N_r enabling optimisation on a reduced number of parameters and thus search space dimensions. The following stage further tunes a subset of the parameters, targeted at the high-current behaviour of the Gummel plots. Finally, all parameters are optimised using the Gummel plots and common-emitter characteristic. This overall optimisation resulted in the final extracted parameter values which are used in the model comparison, shown in Table 4.8.

4.4.4 Results

Comparing values between the Ebers-Moll parameters in Table 4.7 and DC Gummel-Poon parameters in Table 4.8 large discrepancies between the values of the same parameters are found, in particular for β_f and β_r. Invariably this is caused by both the different voltage ranges on which the optimisation is performed on and the different optimisation procedures. What can be assured is that the parameters produce a model that is the closest fit to the measurements as defined by the optimisation procedure, i.e. both sets of parameter values are valid despite not matching.

The first AC128 and second OC44 BJTs were selected for the VA comparison, using the mean final objective value to determine the best fit for both models. Fit to measurements of the optimised Ebers-Moll model is illustrated in Figures 4.17 and 4.18 for the OC44 and AC128 respectively. The selected optimisation ranges of the Gummel plots all demonstrate a good fit to measurements. Limitations of the Ebers-Moll model are highlighted by the deviation from measurements at large current values in each of the Gummel plots, and when $|V_\mathrm{ec}|$ is large in the common-emitter characteristic.

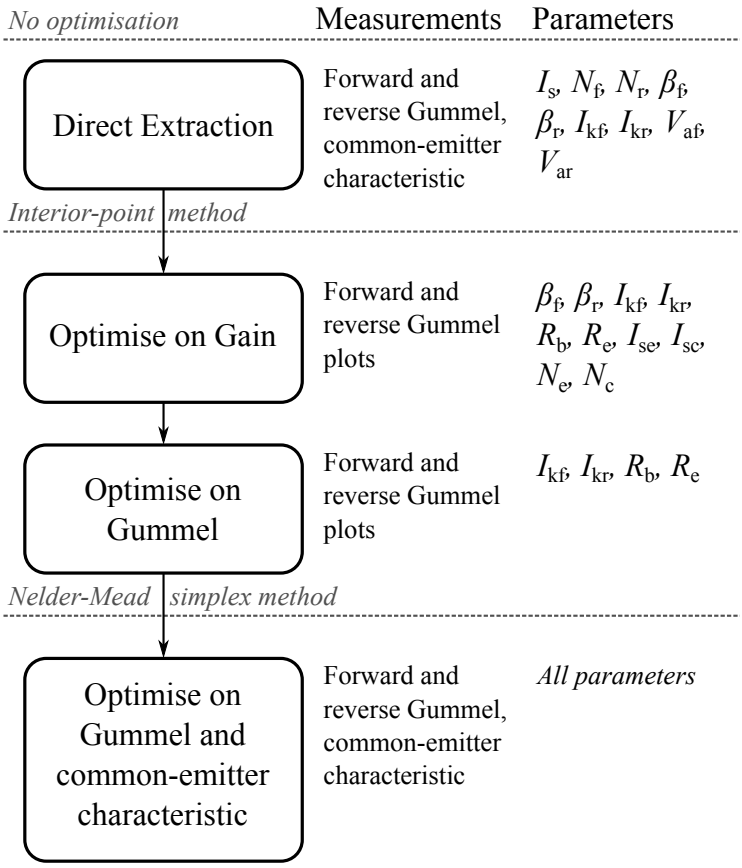


Figure 4.16: The implemented optimisation strategy to find parameter values for the DC Gummel-Poon BJT model.

Table 4.7: Complete set of extracted parameters from all measured germanium BJTs for the Ebers-Moll model. * marks the BJTs used in the comparison.

	AC128				OC44		
	1*	2	3	4	1	2*	3
Obj. Val.	0.1408	0.1454	0.1746	0.1845	0.08730	0.07667	0.8533
Parameter							
I_s (µA)	31.97	30.87	37.39	30.31	2.729	3.150	2.89
β_f (A/A)	51.83	42.53	47.94	43.16	102.0	98.40	214.8
β_r (A/A)	10.12	10.62	7.485	5.036	9.593	12.24	24.49
N_f	1.195	1.142	1.250	1.299	1.127	1.114	1.063
N_r	1.158	1.192	1.229	1.218	1.155	1.137	1.064

Table 4.8: Complete set of extracted parameters from all measured germanium BJTs for the DC Gummel-Poon model.

	AC128				OC44		
	1*	2	3	4	1	2*	3
Obj. Val.	1.038e-3	1.652e-3	1.468e-3	1.466e-3	1.200e-3	0.962e-3	0.822e-3
Parameter							
I_s (µA)	27.77	31.67	29.44	20.66	1.423	2.245	2.867
β_f (A/A)	156.7	125.7	149.9	229.6	307.0	283.5	226.2
β_r (A/A)	54.45	29.18	15.18	14.66	20.27	24.10	19.78
N_f	1.142	1.154	1.171	1.133	1.022	1.067	1.097
N_r	1.135	1.157	1.176	1.140	1.025	1.080	1.110
V_af (V)	19.93	61.24	17.95	19.68	8.167	6.863	8.787
V_ar (V)	54.1	102.3	47.06	88.28	14.84	12.40	15.26
I_kf (mA)	1322	606.2	642.8	463.0	43.82	57.93	109.5
I_kr (mA)	122.5	126.3	411.6	241.5	611.7	1012	477.8
I_se (µA)	3.047	2.485	2.36	2.190	0.03053	0.09406	0.1404
I_sc (µA)	5.920	4.295	5.37	7.545	0.2135	0.3443	0.2958
N_e	2.118	1.836	2.007	1.796	1.316	1.548	2.130
N_c	1.383	1.409	1.372	1.363	1.258	1.494	1.475
R_b (Ω)	2.234	2.403	3.477	1.885	32.83	23.60	15.03
R_e (mΩ)	215.3	183.1	290.2	306.4	968.7	604.9	606.2
R_c (µΩ)	36.15	11.82	7.453	17.27	989.9	22.63	1032

The fit of the DC Gummel-Poon model to the same measurements is illustrated in Figures 4.19 and 4.20 for the OC44 and AC128 respectively. Fit in high current regions of the Gummel plots is markedly improved, and similar improvements can be observed in the common-emitter characteristic.

Despite having several additional terms in the model designed to fit measurements, systematic error is still observed in the optimised DC Gummel-Poon model. This is likely indicative of a combination of two issues: that there are unmodelled effects in the behaviour of the BJT; and that the optimisation strategy landed in a local minimum of the search space. It is also possible that due to the sensitivity of the devices that during the measurement temperature effects were non-negligible leading to distorted measurements. Nonetheless, the fit is clearly improved over that of the Ebers-Moll model and therefore provides an interesting model with which to compare for use in VA models.

4.5 Virtual Analogue comparison of Bipolar Junction Transistor models

Models of both case studies were created using the extracted parameter values. Three metrics were used to compare the differences in BJT models: informal listening tests, a comparison of model output waveforms, and a comparison of computational efficiency. For each test all potentiometers were set to the maximum position, testing different positions caused no substantial difference in results.

Further validation of the BJT model by comparison to circuit measurements were omitted due to the sensitivity of the devices: changes in room temperature could change the behaviour of the BJT such that the circuit behaves significantly different.

4.5.1 Informal listening tests

The focus of the analysis of the models is the objective difference between the outputs of circuit models using the different BJTs. Because of this a formal listening test was excluded, but informal listening tests were performed to get some indication of perceptual differences.

Seven listeners from the field of VA modelling participated, the group consisted of the authors of [97], their colleagues, and attendees of the DAFx 17 conference. The authors and their colleagues used their own personal listening setup which varied between headphones and loudspeakers, while the attendees of DAFx 17 listened through

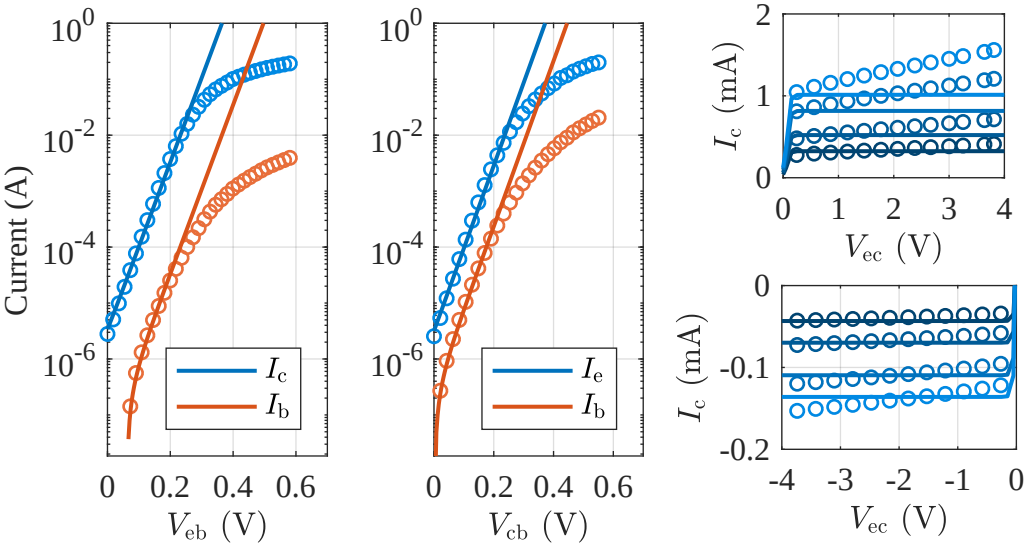


Figure 4.17: Optimised fit of the Ebers-Moll model to measurements of the OC44. Measured points (downsampled by a factor 5 for Gummel plots and 3 for common-emitter) are marked by ∘, lines indicate optimised model. For the common-emitter characteristic, $I_b = 3,\ 5,\ 8,\ 10\,\mu A$.

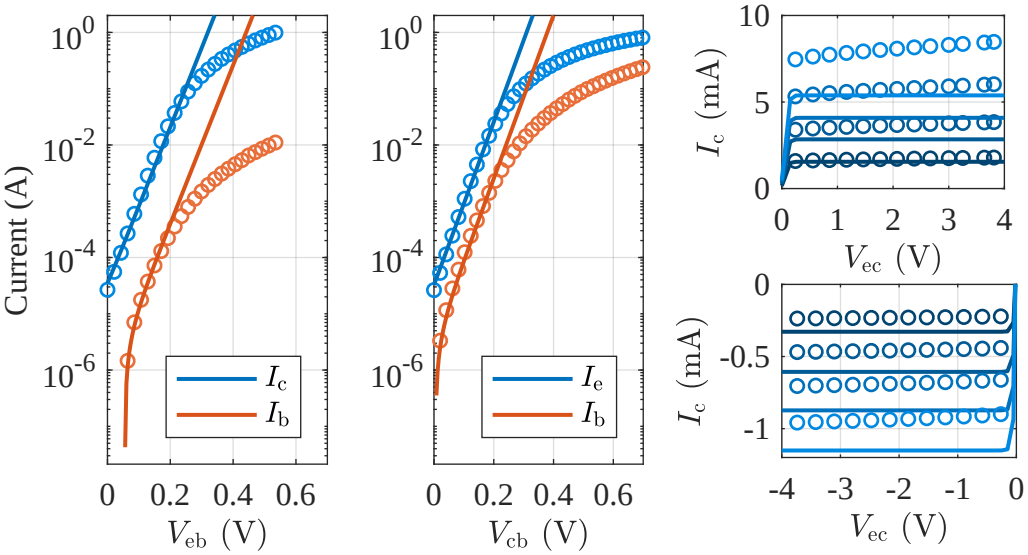


Figure 4.18: Optimised fit of the Ebers-Moll model to measurements of the AC128. Measured points (downsampled by a factor 5 for Gummel plots and 3 for common-emitter) are marked by ∘, lines indicate optimised model. For the common-emitter characteristic, $I_b = 26,\ 51,\ 75,\ 100\,\mu A$.

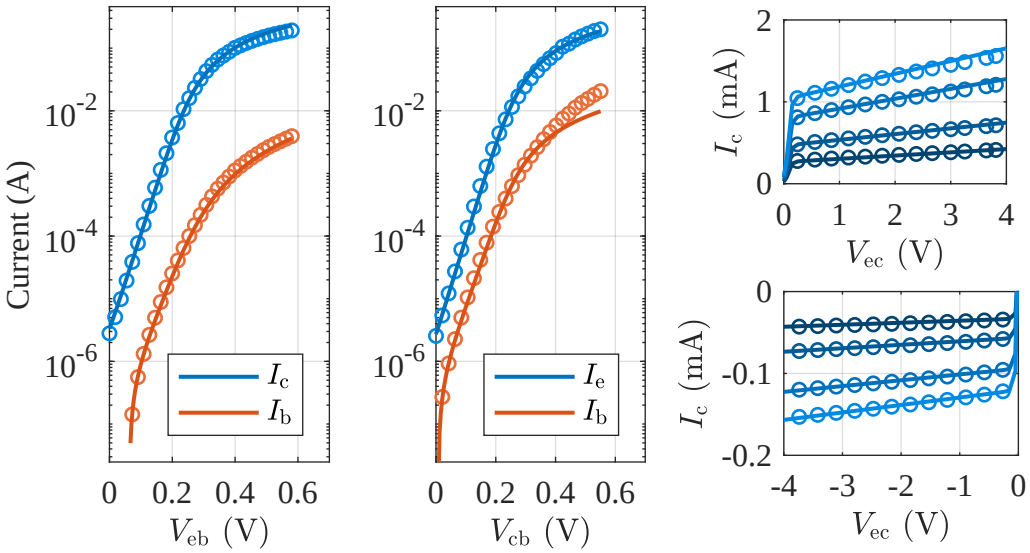


Figure 4.19: Optimised fit of the DC Gummel-Poon model to measurements of the OC44. Measured points (downsampled by a factor 5 for Gummel plots and 3 for common-emitter) are marked by ∘, lines indicate optimised model. For the common-emitter characteristic, $I_b = 3,\ 5,\ 8,\ 10\,\mu A$.

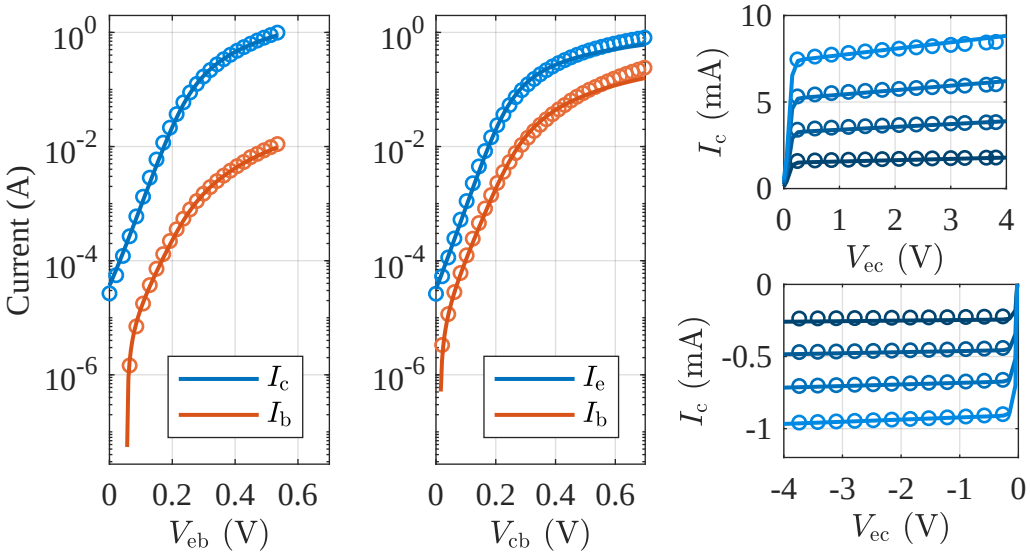


Figure 4.20: Optimised fit of the DC Gummel-Poon model to measurements of the AC128. Measured points (downsampled by a factor 5 for Gummel plots and 3 for common-emitter) are marked by ∘, lines indicate optimised model. For the common-emitter characteristic, $I_b = 26,\ 51,\ 75,\ 100\,\mu A$.

a set of closed-back studio headphones driven by a headphone output of a laptop.

Two guitar signals – a short riff and a single chord – were processed by both case studies at $8\times$ oversampling as a means of comparing each model, available at `https://bholmesqub.github.io/DAFx17/`. One listener from DAFx 17 reported that the difference between samples is minimal. The remaining listeners agreed that differences could be heard between each model, with the Ebers-Moll model having the most high frequency content due to distortion and the AC Gummel-Poon having the least.

Further sound examples are provided for readers[1]. The new sound examples feature cleaner guitar recordings and amp simulations to demonstrate the effect in context.

4.5.2 Waveform comparison

An objective comparison of each BJT model is achieved here using time-domain waveforms. Sinusoids at different frequencies and amplitudes were processed by both case studies and each model. To remove transient behaviour from the results the waveform was repeated for $2\,\mathrm{s}$, with the final period of each of these signals shown in Figure 4.22 and 4.21 for the Rangemaster and Fuzz Face respectively. Plots at $1200\,\mathrm{Hz}$ show the largest difference for the AC effects, illustrating the low-pass type behaviour of the capacitances. Differences due to the increased DC complexity are most prominent at lower amplitudes.

4.5.3 Computational efficiency

To understand the cost of increasing the complexity of the BJT model the computational efficiency of each model was compared. An Ebers-Moll model with C_{eb} and C_{cb} (AC Ebers-Moll) was included for this test to provide an improved assessment of the cost of the capacitances. Three metrics were selected: average time required for one second of simulation, average iterations, and average sub-iterations as Damped Newton's method was chosen to solve the nonlinear equations.

The test drives both case studies with a guitar signal for each BJT model, with the peak amplitude of the signal set to 20 different levels. Computation time was then measured by MATLAB's `cputime` function to measure the simulation time independent from other processes that may use computer system time. The results are shown in Table 4.9. It is clear from the results that increasing the DC complexity causes a significant increase in computation time, whereas including additional capacitances carries

[1] https://bholmesqub.github.io/thesis/chapters/bjt-modelling/

Table 4.9: Mean simulation time required to process one second of signal, iterations per sample, and sub-iterations per sample of circuit models processing a guitar chord using different BJT models. The Rangemaster was tested over a peak voltage range of $0.1 - 2\,\mathrm{V}$, the Fuzz-Face over a range of $10 - 100\,\mathrm{mV}$.

Model	Rangemaster		Fuzz-Face	
	Sim. time/s (ms)	Mean Iter. /Sub-iter.	Sim. time/s (ms)	Mean Iter. /Sub-iter.
DC E.M.	65.3	3.56/0.28	148.3	3.64/0.03
AC E.M.	38.4	3.50/0.19	150.5	3.58/0.01
DC G.P.	207.9	3.53/0.20	421.5	3.03/0.03
AC G.P.	205.2	3.49/0.13	411.2	2.99/0.01

little cost. As iterations and sub-iterations decrease with increasing model complexity the increase in computation must be due to the increased complexity of evaluating the model equations. Decrease in computation cost when including the capacitances can be attributed to the reduction in high frequencies reducing the stress placed on the iterative solver, outweighing the increase in the complexity from including additional components.

4.6 Conclusion

A comparison of BJT models has been presented with a focus on germanium BJTs. A model similar to the Gummel-Poon was described as an extension to the Ebers-Moll model. Both Ebers-Moll and DC Gummel-Poon models were characterised by extracting parameters from measured data using a multi-step optimisation strategy. Models with extracted parameter values demonstrated a good fit to the measurements, with the Ebers-Moll model failing where anticipated, and the DC Gummel-Poon offering improved fit in regions where the Ebers-Moll model exhibited high error.

The resultant models were compared through the use of two case study circuits covering both moderately and highly distorted circuit outputs. The circuit models were compared using three metrics: audible and waveform differences, and computational efficiency. Results show that increase in model complexity does make a change to the behaviour of germanium BJTs in VA circuit models. This work has primarily focused on improving DC characterisation; however, the results show that AC effects are at least equally important. The improved DC characterisation has a significant increase

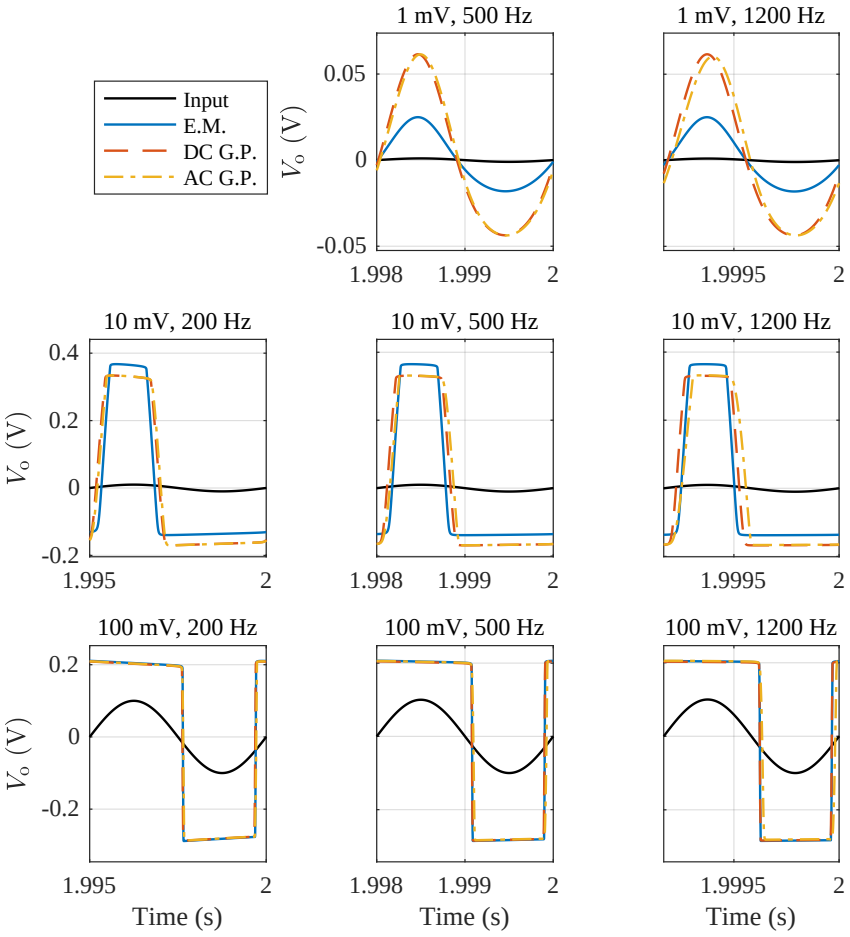


Figure 4.21: Single cycle waveforms of sine waves at various frequencies and amplitudes, and the respective Fuzz Face output with different BJT models.

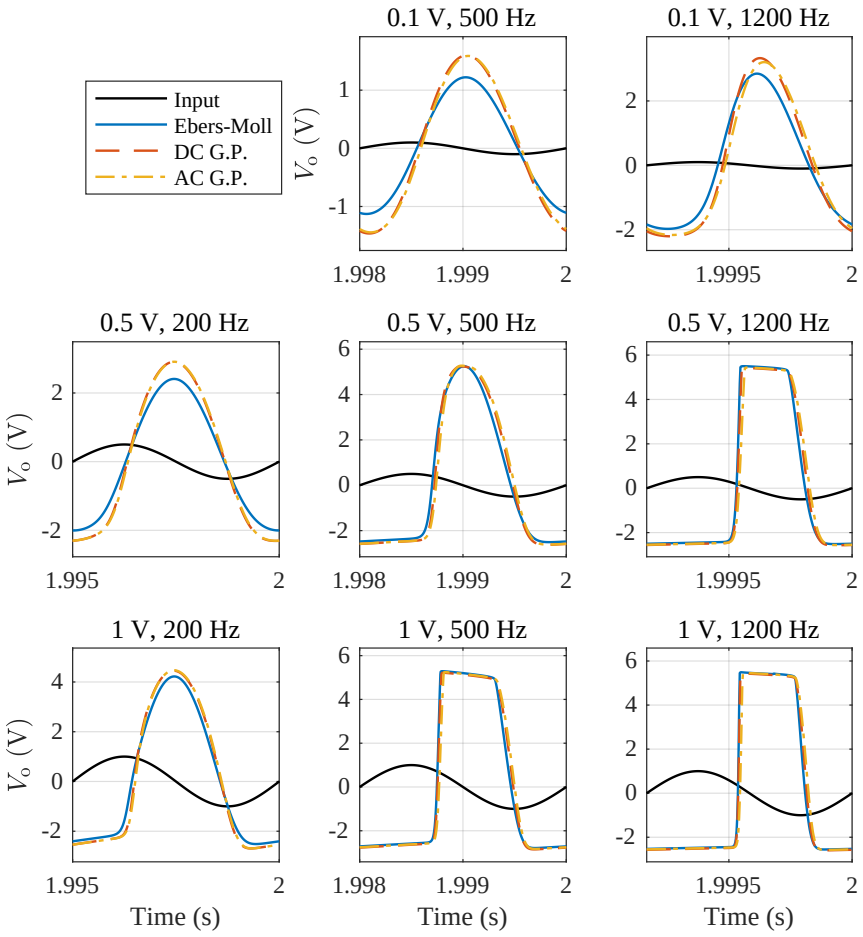


Figure 4.22: Single cycle waveforms of sine waves at various frequencies and amplitudes, and the respective Rangemaster output with different BJT models.

in computational cost whereas the cost of the AC effects are minimal. These results indicate that any first extension to the Ebers-Moll model should be AC effects, and further extensions should then concern DC effects.

The core motivating factor for implementing and characterising more sophisticated BJT models was to reduce the error present in VA circuits featuring germanium BJTs. Should modellers encounter situations where the Ebers-Moll model is insufficient an extended model has been validated which will improve accuracy. An implementation of the Gummel-Poon model has been included in ACME.jl[2] emulation tool for modellers interested in further investigation.

[2] `https://github.com/HSU-ANT/ACME.jl`

Chapter 5

Identification problem design and analysis

A major objective of the work presented in this thesis is to develop strategies for minimising the difference between a circuit and its model, exclusively operating on input/output data. In the following chapter, this objective is divided into two strategies, calibration and parameter estimation, both utilising similar factors though they are applied differently. Prior to their application to challenging case studies in the following chapters, here maximally simple case studies are examined, enhancing the study of each factor of both calibration and parameter estimation and how both identification problems are formed.

The general problem formulation shared by both strategies is illustrated in Figure 5.1. An excitation signal V_{i} drives both the circuit and the model. An objective function then compares the input/output data to yield a value of ξ which enumerates the difference between model and circuit. The values of the physical parameters in $\boldsymbol{\theta}$ define the behaviour of the chosen circuit model and are tuned using an optimisation algorithm to minimise the value of ξ. From this diagram, the individual elements required for a successful identification are indicated: a suitable physical model and a combination of objective function and excitation signal that exposes the desired behaviour of the circuit to be modelled. As mentioned in Chapter 1, the presented identification strategy requires the circuit topology and a set of initial component parameters, typically extracted from a schematic.

Analysis of the objective function, excitation signal, and model is required to determine whether they are sufficient to create a model that matches with the target device. Broadly this depends on the objectives of the identification. Calibration attempts to capture the input/output behaviour of the circuit over a predefined range of operation

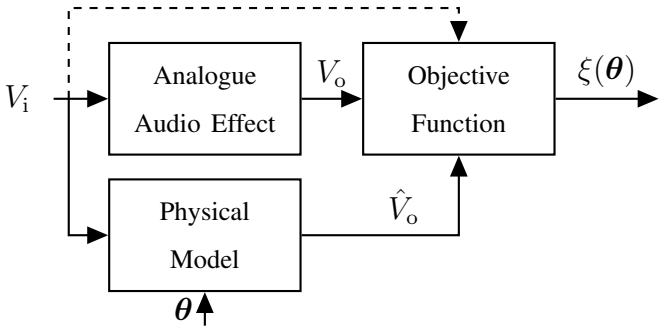


Figure 5.1: Diagram illustrating the evaluation of a parameter set $\boldsymbol{\theta}$ by comparison of a physical model with said parameters and the desired analogue audio effect. The dashed line indicates the use of the input signal in the case that the objective function evaluates the transfer function of the circuit.

which is represented by the objective function, aiming to achieve this with a minimal investment of effort. On the other hand, parameter estimation focusses on retrieving accurate parameter values for the model such that the results can be used in further analysis of the circuit, or extended models using the identified model as a building block, which may demand a larger investment in analysis to provide results that are more widely applicable.

Identifying circuits with the objective of estimating the physical circuit parameters is assumed to be the more challenging of the two strategies, likely requiring a highly accurate underlying model. Symbolic analysis of the model is employed in an attempt to determine whether the parameters are in fact estimable. Whether the parameters are estimable with the selected data from the excitation signal/objective function must then be verified using simulated data.

With the only requirement for calibration being the capture of the circuit behaviour, there is no need to investigate whether the model parameters are estimable. The search is instead for a technique of improving the speed of the identification process which in itself requires minimal computation and user time.

Analyses of both calibration and parameter estimation are performed to determine whether they satisfy their individual objectives, and a rudimentary comparison is performed to highlight the differences between strategies. First suitable models are designed for use in the identification procedure.

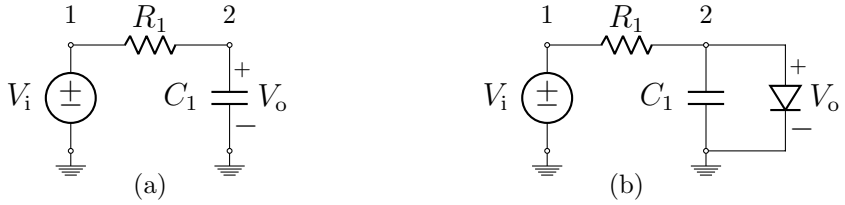


Figure 5.2: (a) A basic RC circuit, (b) The SSDC from Figure 3.1, both with labelled nodes.

5.1 Model design

The case study circuits featured in this chapter are an RC circuit and the single sided diode clipper (SSDC), pictured in Figure 5.2. Though the end objective is to identify nonlinear circuits, by selecting the first case to be a linear circuit simplifies the identification task, and once a suitable method has been found for it, the method can then be extended for nonlinear behaviour.

A schematic is assumed to be available for the circuit being modelled from which the circuit topology can be extracted. Models of the constituent components must then be selected, each component parameter is gathered into the circuit parameter vector $\boldsymbol{\theta}$. Depending on whether the circuit is linear or nonlinear (all presented circuits are time invariant with the exception of potentiometers, which are constrained to time invariance) a suitable modelling paradigm is then chosen and the circuits are turned into computable models. Finally these models can be analysed to determine whether the value of each parameter can be estimated, though this is not essential if the objective is solely the capture of circuit behaviour.

5.1.1 Parameterisation

Prior to the design of a circuit-level model, component models must first be selected from which a set of parameters is defined. For any given component a range of models exists with different levels of complexity and numbers of parameters. This has been seen in the previous chapter where the Ebers-Moll model which has 5 parameters is extended to the AC Gummel-Poon model with 13. Components that are typically represented ideally with linear models may also be extended in scenarios that warrant additional complexities, for example, the resistor exhibits dynamic behaviour at high frequencies [98].

Each parameter included in the model increases the number of dimensions in the

search space of which the optimisation algorithm must find the minimum. It therefore benefits to begin with the simplest component models to reduce the amount of work the optimisation algorithm must perform. A trial-and-error approach must be used to find the optimal set of component models: if the circuit behaviour cannot be fit by its model, some behaviour has been excluded which requires either additional complexity in the components or a change in topology that often involves a change in components.

The simplest component models of a resistor and capacitor are the single parameter component laws given in Chapter 2.1.1. For the RC circuit, this yields the parameter set

$$\boldsymbol{\theta}_{\mathrm{rc}} = [R_1 \quad C_1]. \tag{5.1}$$

The Shockley model is chosen for the SSDC, adding two parameters to the set of the RC circuit,

$$\boldsymbol{\theta}_{\mathrm{dc}} = [R_1 \quad C_1 \quad I_{\mathrm{s}} \quad N]. \tag{5.2}$$

Ideal values and respective tolerance for parameters can be found from the schematic, though for multi-parameter component models, often there are not readily available values. In this situation datasheets of the multi-parameter component can be used to find initial values, or alternatively extracted from a set of measurements on equivalent devices not embedded in the circuit. This chapter only uses simulated data so parameter values do not need to refer to a specific circuit, $R_1 = 2.2\,\mathrm{k\Omega}$, $C_1 = 10\,\mathrm{nF}$, $I_{\mathrm{s}} = 10\,\mathrm{fA}$, $N = 1$.

5.1.2 Modelling paradigm

With a set of component models defined it is then possible to create a circuit-level model. As previously stated in Chapter 2.1.3, MNA is used to derive each model used in this work. Two forms derived from MNA are used for linear and nonlinear models: transfer functions for the linear RC circuit, and a state-space model for the SSDC. Transfer functions offer a method of representing a wide frequency range in a low number of data points in comparison to a time domain signal. Nonlinear behaviour is more challenging to represent in the frequency domain so by choosing a time domain state-space model, the nonlinear behaviour can be properly captured. Derivations of each model are worked through here to demonstrate the link from MNA equations to the final model form.

It is important to note that transfer functions do not directly relate to the diagram in Figure 5.1 which is described in the time domain. In this situation, a Fourier transform (specifically a DFT) is effectively placed in the objective function and the input is

deconvolved from the circuit and model outputs to produce the transfer functions to be compared, though these tasks are performed offline such that they are not repeated during optimisation. Further discussion is provided in Section 5.2.1.

RC circuit

The RC circuit illustrated in Figure 5.2 has 2 nodes and 1 voltage source producing 3 unknowns, the nodal voltages V_1, V_2, and the current through the voltage source I_{vi}. Using MNA this circuit is characterised by

$$\frac{V_1 - V_2}{R_1} + I_{\text{vi}} = 0 \quad \text{(Node 1)}, \tag{5.3}$$

$$-\frac{V_1 - V_2}{R_1} + C_1 \frac{dV_2}{dt} = 0 \quad \text{(Node 2)}, \tag{5.4}$$

$$V_1 = V_{\text{i}} \quad \text{(Source } V_{\text{i}}\text{)}. \tag{5.5}$$

Here the conventions used are current flowing out of a node is seen as positive, and current flows from positive to negative voltages. The nodal equations can be represented in matrix form,

$$\begin{bmatrix} G_{\text{R1}} & -G_{\text{R1}} & 1 \\ -G_{\text{R1}} & G_{\text{R1}} & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ I_{\text{v}} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & C_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1' \\ V_2' \\ I_{\text{v}}' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} V_{\text{i}}. \tag{5.6}$$

As a transfer function is the desired output of the derivation, the voltage derivatives can instead be replaced by the steady-state term of the Laplace transform, i.e. let the derivative become a part of the conductance of the capacitor, in this case $G_{\text{C1}} = j\omega C_1$ such that,

$$\begin{bmatrix} G_{\text{R1}} & -G_{\text{R1}} & 1 \\ -G_{\text{R1}} & G_{\text{R1}} + G_{\text{C1}} & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ I_{\text{v}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} V_{\text{i}}. \tag{5.7}$$

Inverting the matrix and using a selection vector $\mathbf{N}_{\text{O}} = [0\ 1\ 0]$ to define V_2 as the output of the model,

$$V_2 = \mathbf{N}_{\text{O}} \begin{bmatrix} V_1 \\ V_2 \\ I_{\text{v}} \end{bmatrix} = \mathbf{N}_{\text{O}} \begin{bmatrix} G_{\text{R1}} & -G_{\text{R1}} & 1 \\ -G_{\text{R1}} & G_{\text{R1}} + G_{\text{C1}} & 0 \\ 1 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} V_{\text{i}}, \tag{5.8}$$

which can finally be rewritten as the transfer function of the circuit:

$$H(s) = \frac{V_2}{V_{\text{i}}} = \frac{1}{1 + sR_1C_1}. \tag{5.9}$$

This equation can now be used to compute the frequency response of the RC circuit by specifying values of ω and letting $s = j\omega$.

Single sided diode clipper

As the SSDC is simply the RC circuit with an additional diode across the output terminals, the nodal equations are similar but with the added diode VCCS term $I_{\mathrm{d}}(V_2)$,

$$\frac{V_1 - V_2}{R_1} + I_{\mathrm{vi}} = 0 \qquad \text{(Node 1)} \tag{5.10}$$

$$-\frac{V_1 - V_2}{R_1} + C_1 \frac{dV_2}{dt} = I_{\mathrm{d}}(V_2) \quad \text{(Node 2)} \tag{5.11}$$

$$V_1 = V_{\mathrm{i}}. \qquad \text{(Source } V_{\mathrm{i}}\text{)} \tag{5.12}$$

Again this can be expressed using matrix notation, with the dynamic and static parts separated, i.e. capacitances operate on the nodal voltage derivative vector,

$$\begin{bmatrix} G_{\mathrm{R1}} & -G_{\mathrm{R1}} & 1 \\ -G_{\mathrm{R1}} & G_{\mathrm{R1}} & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ I_{\mathrm{v}} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & C_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1' \\ V_2' \\ I_{\mathrm{v}}' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} V_{\mathrm{i}} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} I_{\mathrm{d}}(V_2) \tag{5.13}$$

To derive a state-space model from this form it must first be discretised as described in Section 2.1.3. A state term x_{c} is introduced and the conductance of the capacitor is redefined as $G_{\mathrm{C1}} = 2f_{\mathrm{s}}C_1$, with the state updated by $x_{\mathrm{c}}(n) = 2G_{\mathrm{C1}}V_2(n) - x_{\mathrm{c}}(n-1)$. The remaining part of the discrete model is given by

$$\begin{bmatrix} G_{\mathrm{R1}} & -G_{\mathrm{R1}} & 1 \\ -G_{\mathrm{R1}} & G_{\mathrm{R1}} + G_{\mathrm{C1}} & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1(n) \\ V_2(n) \\ I_{\mathrm{v}}(n) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} x_{\mathrm{c}}(n-1) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} V_{\mathrm{i}}(n) + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} I_{\mathrm{d}}(V_2(n)). \tag{5.14}$$

From (5.14), the necessary incidence matrices can be found from the model,

$$\mathbf{N}_{\mathrm{O}} = \mathbf{N}_{\mathrm{C}} = \mathbf{N}_{\mathrm{N}} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \tag{5.15}$$

so that the state-space matrices can be derived as in Chapter 2.1.2,

$$\mathbf{A} = \frac{4C_1R_1f_{\mathrm{s}}}{2C_1R_1f_{\mathrm{s}} + 1} - 1 \quad \mathbf{B} = \frac{4C_1f_{\mathrm{s}}}{2C_1R_1f_{\mathrm{s}} + 1} \quad \mathbf{C} = -\frac{4C_1R_1f_{\mathrm{s}}}{2C_1R_1f_{\mathrm{s}} + 1}, \tag{5.16}$$

$$\mathbf{D}_{\mathrm{O}} = \frac{R_1}{2C_1R_1f_{\mathrm{s}} + 1} \qquad \mathbf{E}_{\mathrm{O}} = \frac{1}{2C_1R_1f_{\mathrm{s}} + 1} \quad \mathbf{F}_{\mathrm{O}} = -\frac{R_1}{2C_1R_1f_{\mathrm{s}} + 1}, \tag{5.17}$$

$$\mathbf{D}_{\mathrm{N}} = \mathbf{D}_{\mathrm{O}} \qquad \mathbf{E}_{\mathrm{N}} = \mathbf{E}_{\mathrm{O}} \qquad \mathbf{F}_{\mathrm{N}} = \mathbf{F}_{\mathrm{O}}. \tag{5.18}$$

Specifying a value for f_{s}, the model can now be driven in an identical manner as the circuit, supplying a signal for V_{i} and recording the output V_{o}.

5.1.3 Redundancy analysis and models with estimable parameters

In the design of a model, it is not guaranteed that each parameter is estimable. First consider the case of two resistors in series: basic circuit theory (specifically working from Kirchoff's Current Law) dictates that the total resistance is the sum of the two individual resistances. For example when applied to the RC circuit, the resistance could be shown as two resistances, e.g. let $R_1 = R_{1,1} + R_{1,2}$, resulting in the transfer function

$$H(s) = \frac{V_2}{V_\mathrm{i}} = \frac{1}{1 + s(R_{1,1} + R_{1,2})C_1}. \tag{5.19}$$

If no other components are connected in parallel then the resistance can be summed to one value; equivalent behaviour is observed so long as the sum of the two is the same regardless of their individual values, meaning that they cannot be estimated. This is referred to as parameter redundancy. Astute readers will notice that the parameters only appear in one term as a sum and product meaning that only one parameter is required to control the behaviour of the circuit. Not every case may be as easy to identify as the presented case, and therefore it serves to have a method of analysis to determine whether a model exhibits parameter redundancy.

Originally designed for biological models, here the method from [34] is adapted for circuit models. To identify parameter redundancy, first an exhaustive summary $\boldsymbol{\kappa}$ must be defined that extracts all terms including parameters while excluding other behaviour, e.g. time, voltage, etc. A Jacobian matrix is then found as the partial derivatives $\mathbf{J}(\boldsymbol{\kappa}) = \partial\kappa_i/\partial\theta_j$. By definition $\mathrm{rank}(\mathbf{J}(\boldsymbol{\kappa})) \leq \dim(\boldsymbol{\theta})$ where $\dim(\boldsymbol{\theta})$ is the number of parameters. If the rank of this matrix is found to be lower than the number of parameters i.e $\mathrm{rank}(\mathbf{J}(\boldsymbol{\kappa})) < \dim(\boldsymbol{\theta})$, the model is said to exhibit parameter redundancy, and if $\mathrm{rank}(\mathbf{J}(\boldsymbol{\kappa})) = \dim(\boldsymbol{\theta})$ then each parameter is estimable. Following the analysis and correction of the presented models using this method, whether a model has estimable parameters is demonstrated numerically in Section 5.3.2.

The success of this method is dependent upon determining an accurate exhaustive summary. Many models require some manipulation before they can produce this. Revisiting the case of two series resistors, if represented in MNA form, a section of $\mathbf{S}$ may resemble

$$\begin{bmatrix} G_{\mathrm{R1},1} & -G_{\mathrm{R1},1} & 0 \\ -G_{\mathrm{R1},1} & G_{\mathrm{R1},2} - G_{\mathrm{R1},1} & -G_{\mathrm{R1},2} \\ 0 & -G_{\mathrm{R1},2} & G_{\mathrm{R1},2} \end{bmatrix}, \tag{5.20}$$

which would have a corresponding parameter vector $\boldsymbol{\theta} = [R_{1,1},\ R_{1,2}]$ and exhaustive summary $\boldsymbol{\kappa} = [G_{\mathrm{R1},1},\ G_{\mathrm{R1},2} - G_{\mathrm{R1},1},\ G_{\mathrm{R1},2}]$. Demonstrated in this exhaustive sum-

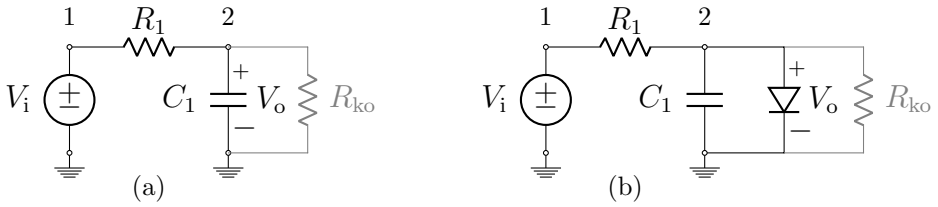


Figure 5.3: Case study circuits adapted to remove redundancy with a known component R_ko. (a) A basic RC circuit, (b) The SSDC from Figure 3.1.

mary, elements which are equal to another element, or equivalent except for the sign, have been excluded as they do not increase the rank of the resulting Jacobian matrix, only increasing the computational expense. To visualise the rank of the Jacobian, the reduced row echelon form can be used as the rank is maintained but the representation is maximally simplified,

$$\mathrm{rank}(\mathbf{J}(\boldsymbol{\kappa})) = \mathrm{rank}\left(\begin{bmatrix} -G_{\mathrm{R1},1}^2 & 0 \\ G_{\mathrm{R1},1}^2 & -G_{\mathrm{R1},2}^2 \\ 0 & -G_{\mathrm{R1},2}^2 \end{bmatrix}\right) = \mathrm{rank}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = 2. \tag{5.21}$$

This exhaustive summary would fail in demonstrating the parameter redundancy as each node is effectively treated as a separate output, but by inspecting the circuit directly it would be clear that the resistances must be combined. This prompts an initial step of inspecting circuit schematics for components in parallel and series which otherwise could be replaced by fewer components and therefore parameters.

RC circuit

As previously mentioned, redundancy in the transfer function of the RC circuit is clear without the application of the symbolic analysis, extracting terms featuring parameters from (5.9) produces the exhaustive summary $\boldsymbol{\kappa}(\boldsymbol{\theta}_\mathrm{rc}) = R_1\, C_1$. The exhaustive summary contains only 1 entry, and therefore it is impossible for the rank of the Jacobian to be equal to the number of parameters, i.e. 2. So long as the product R_1C_1 is the same, different values of R_1 and C_1 will produce models with equivalent behaviour.

Should the RC circuit transfer function be expressed differently, it may not be possible to identify the parameter redundancy in the model. For example by simply dividing the numerator and denominator of the transfer function by R_1,

$$H(s) = \frac{1/R_1}{1/R_1 + sC_1}, \tag{5.22}$$

the exhaustive summary becomes $\boldsymbol{\kappa}(\boldsymbol{\theta}_\mathrm{rc}) = [1/R_1,\ C_1]$, of which the Jacobian has rank equal to the number of parameters. This does not mean that the parameter redundancy

has been removed, but masked. A second fundamental step in this method is ensuring that the number of terms featuring parameters is minimised prior to analysis.

When faced with parameter redundancy, additional information must be added to the model. One such method is to include a component with known behaviour across the output. The RC circuit has been extended with a known output resistance R_{ko} placed across the output terminals, illustrated in Figure 5.3 (a). Following the same derivation from MNA as used to produce (5.9), the extended circuit can be modelled as

$$H(s) = \frac{1}{R_1/R_{\text{ko}} + sR_1C_1}. \tag{5.23}$$

No fewer terms with parameters can be found through model manipulation, so the exhaustive summary is defined as $\boldsymbol{\kappa}(\boldsymbol{\theta}_{\text{rc}}) = [R_1/R_{\text{ko}},\ R_1C_1]$. Examining $\text{rank}(\mathbf{J}(\boldsymbol{\kappa}))$ it is found that the extended model exhibits no parameter redundancy. Intuitively this can be understood as R_1 now also forms a resistor divider, so changes in its value will not only change the product R_1C_1 (which could be compensated for with C_1) but will also change the static output amplitude.

Note that, in principle, components other than a resistor can be used, but if a capacitor were placed across the output port it would sum with the capacitor already present, failing to remove parameter redundancy.

A worked example of the parameter redundancy analysis of the RC circuit has been written in MATLAB, and is available online[1] to further clarify the process.

Single sided diode clipper

Further complications with the method are illustrated in the case of the SSDC. An initial exhaustive summary can be found in the state-space matrices and their combination with the nonlinear component model $\boldsymbol{\kappa}(\boldsymbol{\theta}_{\text{dc}}) = [\mathbf{A},\ \mathbf{B},\ I_{\text{s}}\mathbf{C},\ \mathbf{D}_{\text{O}},\ \mathbf{E}_{\text{O}},\ I_{\text{s}}\mathbf{F}_{\text{O}},\ 1/N]$ where repeated matrices have been omitted as they would not change the rank of the resulting Jacobian matrix, which is found equal to the number of parameters indicating no parameter redundancy. This result is false, but the state-space matrices are complex and do not clearly indicate redundancy, as may be expected from circuits more complex than the RC circuit.

Instead an equivalent model can be used: the parameter focussed form described in Chapter 2.1.3. The current through the voltage source I_{vi} is not used in the computation of the state-space model, and the input voltage source can be directly substituted into

[1] https://bholmesqub.github.io/thesis/chapters/identification-design/redundancy-analysis/

the relevant nodal voltage, such that

$$\begin{bmatrix} G_{\mathrm{R1}} & -G_{\mathrm{R1}} \\ -G_{\mathrm{R1}} & G_{\mathrm{R1}} \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}} \\ V_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & C_1 \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}}' \\ V_2' \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} I_{\mathrm{s}} \left(\mathrm{e}^{\frac{V_2}{N V_{\mathrm{t}}}} - 1 \right). \tag{5.24}$$

Inspecting this form it is found that several terms with parameters may be removed, for example by multiplying by R_1,

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}} \\ V_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & R_1 C_1 \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}}' \\ V_2' \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} R_1 I_{\mathrm{s}} \left(\mathrm{e}^{\frac{V_2}{N V_{\mathrm{t}}}} - 1 \right). \tag{5.25}$$

Now there are only three terms with parameters, $\boldsymbol{\kappa}(\boldsymbol{\theta}_{\mathrm{dc}}) = [R_1 C_1,\ R_1 I_{\mathrm{s}},\ 1/N]$. It is again clear that as the number of elements in $\boldsymbol{\kappa}$ is fewer than the number of parameters in $\boldsymbol{\theta}_{\mathrm{dc}}$, it is impossible for parameters to be directly estimated.

To observe this in the state-space model the parameters must be combined prior to the model derivation, i.e. R_1 becomes 1, C_1 becomes $R_1 C_1$, and I_{s} becomes $R_1 I_{\mathrm{s}}$:

$$\mathbf{A} = \frac{4 C_1 R_1 f_{\mathrm{s}}}{2 C_1 R_1 f_{\mathrm{s}} + 1} - 1 \quad \mathbf{B} = \frac{4 C_1 R_1 f_{\mathrm{s}}}{2 C_1 R_1 f_{\mathrm{s}} + 1} \quad \mathbf{C} = -\frac{4 C_1 R_1 f_{\mathrm{s}}}{2 C_1 R_1 f_{\mathrm{s}} + 1}, \tag{5.26}$$

$$\mathbf{D}_{\mathrm{O}} = \frac{1}{2 C_1 R_1 f_{\mathrm{s}} + 1} \qquad \mathbf{E}_{\mathrm{O}} = \frac{1}{2 C_1 R_1 f_{\mathrm{s}} + 1} \quad \mathbf{F}_{\mathrm{O}} = -\frac{1}{2 C_1 R_1 f_{\mathrm{s}} + 1}, \tag{5.27}$$

$$\mathbf{D}_{\mathrm{N}} = \mathbf{D}_{\mathrm{O}} \qquad \mathbf{E}_{\mathrm{N}} = \mathbf{E}_{\mathrm{O}} \qquad \mathbf{F}_{\mathrm{N}} = \mathbf{F}_{\mathrm{O}}. \tag{5.28}$$

Comparing the above state space matrices with the original matrices (5.16-5.18), they are found to be equivalent. In the output equation, the new values of $\mathbf{D}_{\mathrm{O}}$ and $\mathbf{F}_{\mathrm{O}}$ have been divided by a factor R_1 which is compensated for by a multiplication of R_1 in $\mathbf{B}$ and the diode model. Although the models are equivalent, performing the redundancy analysis on the new exhaustive summary $\boldsymbol{\kappa}(\boldsymbol{\theta}_{\mathrm{dc}}) = [\mathbf{A},\ \mathbf{B},\ R_1 I_{\mathrm{s}} \mathbf{C},\ \mathbf{D}_{\mathrm{O}},\ \mathbf{E}_{\mathrm{O}},\ R_1 I_{\mathrm{s}} \mathbf{F}_{\mathrm{O}},\ 1/N]$ reveals that there is now 1 redundant parameter.

Redundancy exhibited by the SSDC can be removed using the exact same method as with the RC circuit, Figure 5.3 shows the circuit extended with an additional known output resistance R_{ko}. To confirm that the new model has estimable parameters, the parameter focussed representation can again be analysed,

$$\begin{bmatrix} G_{\mathrm{R1}} & -G_{\mathrm{R1}} \\ -G_{\mathrm{R1}} & G_{\mathrm{R1}} + G_{\mathrm{Rko}} \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}} \\ V_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & C_1 \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}}' \\ V_2' \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} I_{\mathrm{s}} \left(\mathrm{e}^{\frac{V_2}{N V_{\mathrm{t}}}} - 1 \right). \tag{5.29}$$

Should the model be multiplied by a factor of R_1, it is now the case that R_1 would still appear independently of the other parameters when it combines with R_{ko}, i.e. $\boldsymbol{\kappa}(\boldsymbol{\theta}_{\mathrm{dc}}) = [R_1/R_{\mathrm{ko}},\ R_1 C_1,\ R_1 I_{\mathrm{s}},\ 1/N]$, from which it is found that $\mathrm{rank}(\mathbf{J}(\boldsymbol{\kappa})) = 4$ and so all parameters are estimable. The same can be found from the state-space

model, defining the common denominator as $G_{\mathrm{D}} = 1/(2C_1R_1f_{\mathrm{s}} + R_1 + R_{\mathrm{ko}})$, the new coefficients are given by

$$\mathbf{A} = G_{\mathrm{D}}4C_1R_1R_{\mathrm{ko}}f_{\mathrm{s}} - 1 \quad \mathbf{B} = G_{\mathrm{D}}4C_1R_1R_{\mathrm{ko}}f_{\mathrm{s}} \quad \mathbf{C} = -G_{\mathrm{D}}4C_1R_1R_{\mathrm{ko}}f_{\mathrm{s}}, \tag{5.30}$$

$$\mathbf{D}_{\mathrm{O}} = G_{\mathrm{D}}R_{\mathrm{ko}} \qquad \mathbf{E}_{\mathrm{O}} = G_{\mathrm{D}}R_{\mathrm{ko}} \qquad \mathbf{F}_{\mathrm{O}} = -G_{\mathrm{D}}R_{\mathrm{ko}}, \tag{5.31}$$

$$\mathbf{D}_{\mathrm{N}} = \mathbf{D}_{\mathrm{O}} \qquad \mathbf{E}_{\mathrm{N}} = \mathbf{E}_{\mathrm{O}} \qquad \mathbf{F}_{\mathrm{N}} = \mathbf{F}_{\mathrm{O}}. \tag{5.32}$$

An important property of the SSDC to note is that the parameter N is always estimable due to its encapsulation in the exponent of the nonlinear function. No matter how the parameters have been combined, N does not directly combine as it is a coefficient of a voltage that appears as an index, whereas the other parameters appear at base level. This will be further investigated numerically in Section 5.3.2.

Redundancy analysis summary

To recap, to find whether the parameters are estimable for a given circuit model, the presented process is:

1. Ensure there is no duplicate behaviour in the circuit itself, e.g. parallel or series components without other connections.

2. Reduce the model to its simplest form, i.e. containing the least number of terms featuring parameters.

3. Apply symbolic redundancy analysis by collecting the terms featuring parameters into an exhaustive summary, and finding the rank of the Jacobian of this vector with respect to the parameter vector.

The parameter focussed version of MNA is used in conjunction with this process to aid particularly in step 2 where it can be difficult to determine whether the model is maximally simplified if using a more condensed modelling paradigm.

Should redundancy be detected, further information must be included into the model. Here an additional component was included with known behaviour, but other methods can also be applied, for example adding an additional output like I_{vi}, though demonstrating this is outside of the scope of this work.

5.2 Formulation of the optimisation problem

With a suitable model, the remaining elements in designing a successful optimisation problem are the excitation signal and objective function. These elements are complementary; the objective function can only enumerate differences between circuit and model that is exposed by the excitation signal.

5.2.1 Excitation signal

The term 'excitation signal' here describes that which is used to drive the signal input of the device under test, V_{i}. Some of the circuits presented also require a voltage to power the circuit, V_{c}, and while this voltage is important in controlling the behaviour of the circuit, V_{c} is not treated as a second signal input but instead fixed at a DC voltage.

To expose both dynamic behaviour the excitation signal must contain a range of frequencies. As mentioned in Chapter 2.2, a common strategy for this is a swept-sine signal. Nonlinear audio circuits also require a range of amplitudes to comprehensively expose the circuit behaviour, which if using a swept-sine would require multiple repeats of the signal at different amplitudes. Instead a multi-sine signal is selected, consisting of a sum of sinusoids between two frequency boundaries and modulated by an amplitude window $w(n)$,

$$V_{\mathrm{i}}(nT) = V_{\mathrm{p}} \cdot w(n) \sum_{d=d_{\mathrm{l}}}^{d_{\mathrm{u}}} A_d \cos(2\pi\ d\ f_0\ nT\ +\ \phi_d), \quad n = 0, \ldots, N_{\mathrm{s}} - 1 \tag{5.33}$$

where V_{p} is the peak voltage, A_d and ϕ_d are the amplitude and phase of sine component d. The distance between sine wave component frequencies is dictated by $f_0 = f_{\mathrm{s}}/N_{\mathrm{s}}$ such that each sine component will have an integer number of periods in the signal. The lower and upper boundaries d_{l} and d_{u} provide a method of bandlimiting the signal, by selecting values closest to the desired lower and upper frequency boundaries. Bandlimiting is a desirable property as it enables a convenient method of focusing the measurements, for example, to limit the amplitude of distortion product frequencies that exceed the Nyquist frequency.

Using the values of A_d to weight the input signal frequency components can help to improve the measured SNR over the desired frequency range and place a focus on different frequencies when the objective function directly operates on V_{o}. Figure 5.4 shows the amplitude response of the RC circuit and the inverse transfer function normalised such that the maximum value is 0 dB. By sampling the inverse of the modelled amplitude response to find values for A_d, the corresponding output of the RC circuit

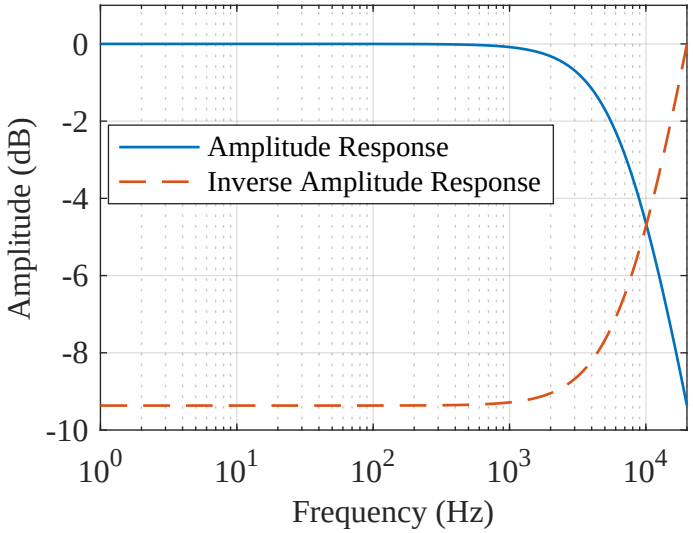


Figure 5.4: Amplitude and inverse amplitude response of the RC circuit. Responses have been normalised such that their maximum value is at 0 dB.

will have an approximately flat frequency response, though due to error in the initial parameters, some mismatch is inherent. Assuming that the noise of the circuit and measurement equipment is Gaussian white noise, this would produce an markedly improved SNR across the measured frequency range. In real measurement scenarios, this is unlikely to be the case, though the values of A_d can be adapted to compensate for noise with different properties should the initial measured data be insufficient to enable a successful identification.

Phase values ϕ_d are generated using Schroeder phases [99],

$$\phi_d = -2\pi \sum_{l=1}^{d-1} (d-l) A_d, \quad d = d_\mathrm{l}, d_\mathrm{l}+1, ..., d_\mathrm{u}. \tag{5.34}$$

This selection of phases distributes the sinusoids to try to minimise the peak to peak voltage, creating a multi-sine signal with low crest-factor. For the generated values of ϕ_d to produce a low crest-factor multi-sine signal values of A_d must satisfy the equation

$$\sum_{d=d_\mathrm{l}}^{d_\mathrm{h}} A_d = 1. \tag{5.35}$$

A flat amplitude envelope maximises the input SNR over the time domain of the signal. Illustrated in Figure 5.5 is a comparison between all phase terms $\phi = 0$ and Schroeder phases. When $\phi = 0$ the amplitude of the middle of the signal is low and likely to be masked by noise in the circuit.

When working with a system that is assumed to be linear, the input signal can be deconvolved from the measured output data, resulting in the measured transfer function. This can then be directly compared to the modelled transfer function evaluated at frequencies given by $d\ f_0$ where $d = d_\mathrm{l}, ..., d_\mathrm{u}$. In this case a flat amplitude window is chosen to prevent adding complexity in the finding of the transfer function, i.e. $w(n) = 1,\ \ n = 1, \ldots, N_\mathrm{s} - 1$. To avoid discontinuities at the start and end of the signal caused by the Schroeder phases, the zero-crossing with the minimum difference between neighbouring samples is chosen to be the point at which the signal starts.

For nonlinear systems, it is beneficial for the excitation signal to vary over a range of amplitudes to capture the amplitude-dependence of the nonlinearity. A Hann window is applied to achieve the range of amplitudes, defined by

$$w(n) = \frac{1}{2}\left(1 - \cos\left(\frac{2\pi n}{N_\mathrm{s} - 1}\right)\right). \tag{5.36}$$

This also leaves the system close to a point of equilibrium at the end of the signal to ensure that multiple periods of the excitation signal are the same, which is necessary for averaging of measurements.

Finally, the excitation signal is scaled such that the peak voltage is at a specified value V_p. A complete implementation of the signal described in this section is available in the form of MATLAB code [2].

RC circuit

The RC circuit is assumed to be linear, and as such there is no need for an amplitude window. The frequency range of interest is the commonly referred to audio band of $20\,\mathrm{Hz} - 20\,\mathrm{kHz}$. To capture at least one cycle of $20\,\mathrm{Hz}$ the signal length must be $50\,\mathrm{ms}$. Selecting the sample rate to be $f_\mathrm{s} = 48\,\mathrm{kHz}$, $N_\mathrm{s} = 2400$, and therefore $f_0 = 20\,\mathrm{Hz}$, requiring $d_\mathrm{l} = 1$ and $d_\mathrm{h} = 1000$. Taking the DFT of the output and deconvolving the input results in a transfer function with 1000 points between $20\,\mathrm{Hz} - 20\,\mathrm{kHz}$. As these cases operate only on simulated data, it is not necessary to apply an inverse frequency response to the signal as the SNR of floating point arithmetic is over what can be expected of even the best measurement setups.

The specified signal is only an example of what could be used to drive a circuit to produce the desired transfer function. As in this chapter only simulated data is used, the transfer function is directly evaluated at 1000 points spaced linearly over the frequency range, drastically reducing computational requirements.

[2] https://bholmesqub.github.io/thesis/chapters/identification-design/multi-sine/

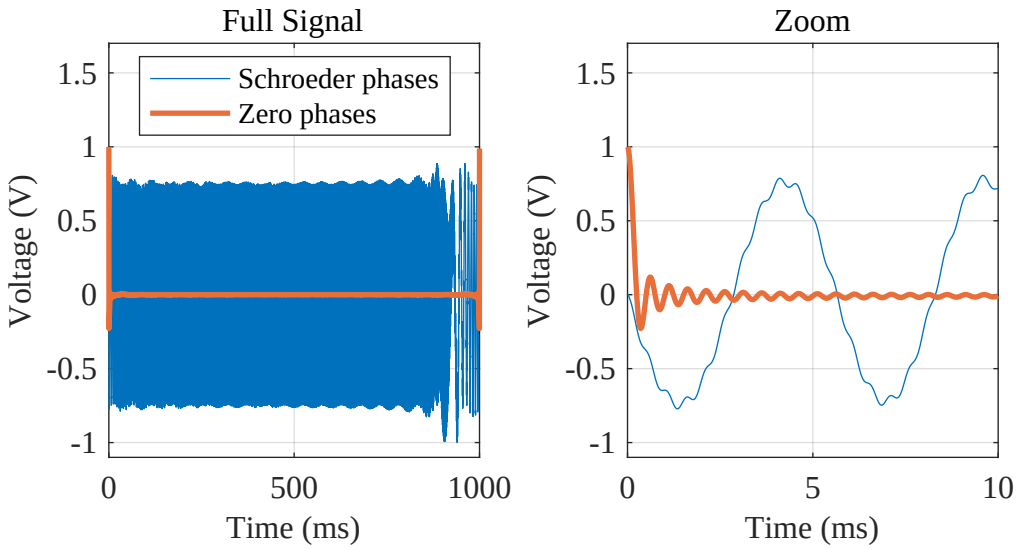


Figure 5.5: Constructive interference of a multi-sine signal with $\phi = 0$ for each component, and Schroeder phases for comparison. Both multi-sine signals are of $1\,\mathrm{s}$ duration, $48\,\mathrm{kHz}$ sample rate, feature components between $20\,\mathrm{Hz}$ and $2\,\mathrm{kHz}$ and are normalised to $1\,\mathrm{V}$ peak.

Data sets consisting of the transfer functions are simulated using the models of the RC circuit with and without the known component $R_{\mathrm{ko}} = 300\,\Omega$, their amplitude responses illustrated in Figure 5.6.

Single sided diode clipper

Distortion of nonlinear circuits typically requires oversampling to reduce aliasing, but increasing the sampling frequency also increases the number of data points which increases the time the optimisation algorithm requires. For this reason the sample rate was set at a common, not oversampled, audio sample rate, $f_{\mathrm{s}} = 48\,\mathrm{kHz}$, and instead the upper frequency limit was chosen to be $8\,\mathrm{kHz}$ such that the cutoff frequency of the RC circuit was contained, but frequencies close to the Nyquist frequency were omitted. To reduce the computation time required to simulate the model driven by the excitation signal, the signal length was set as $5\,\mathrm{ms}$, resulting in a value of $f_0 = 200\,\mathrm{Hz}$. The length of the signal in samples is therefore $N_{\mathrm{s}} = 240$, requiring $d_{\mathrm{l}} = 1$ and $d_{\mathrm{h}} = 40$ to produce the desired frequency limits.

The resulting input signal and outputs of both the SSDC and the circuit extended with the known component $R_{\mathrm{ko}} = 30\,\mathrm{k}\Omega$ are illustrated in Figure 5.7.

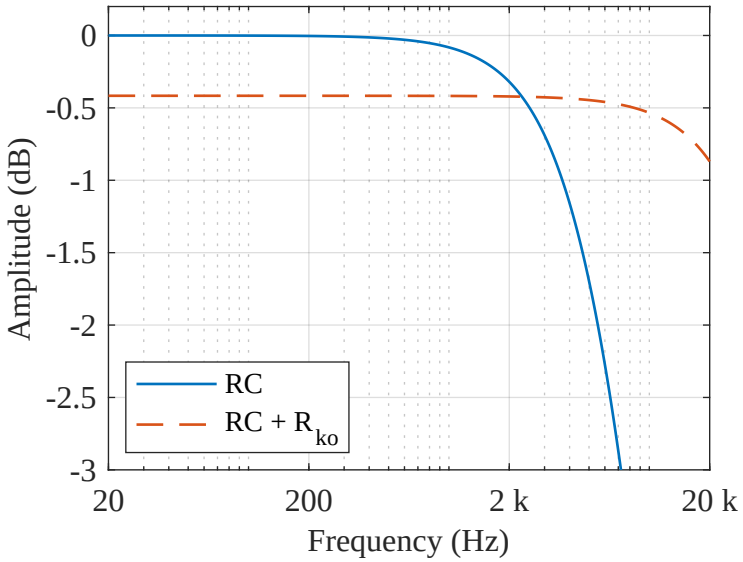


Figure 5.6: Amplitude response of the RC circuit transfer functions used for optimisation, both with and without $R_{\text{ko}} = 300\,\Omega$.

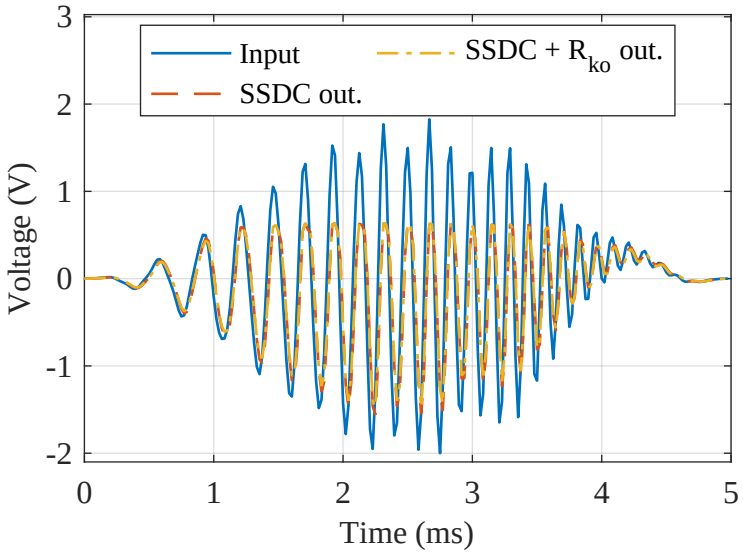


Figure 5.7: Time domain input/output signals used for optimising the single-sided diode clipper.

5.2.2 Objective function

Objective functions are designed to provide a scalar error value $\xi(\hat{\boldsymbol{\theta}})$ that represents the magnitude of difference between measured and simulated data for the current estimated set of parameter values $\hat{\boldsymbol{\theta}}$. A common choice for objective functions is to use a least-squares metric, i.e. minimising the square of the error, which has been applied in VA literature, e.g. [13]. For the current estimated values of the parameter set $\hat{\boldsymbol{\theta}}$, a least-squares objective function is given by

$$\xi(\hat{\boldsymbol{\theta}}) = \frac{1}{\eta} \sum_{n=0}^{N_\mathrm{s}-1} \epsilon(n, \hat{\boldsymbol{\theta}})^2. \tag{5.37}$$

Here n is used as a general index term for evaluating both discrete frequency and time domain data. For both quantities, ϵ is defined equivalently,

$$\epsilon(n, \hat{\boldsymbol{\theta}}) = y(n) - \hat{y}(n, \hat{\boldsymbol{\theta}}), \tag{5.38}$$

where y is the measured signal and $\hat{y}$ is the estimated modelled signal, and y can represent either $H(s_n)$ or $V_\mathrm{o}(nT)$. The discrete index, n, is used to move through both outputs: nT representing the current time step, and $s_n = j\omega_n$ the current frequency. Normalisation factor is given by $\eta = \sum_{n=0}^{N_\mathrm{s}} y(n)^2$ which is proportional to the energy in the signal, and was chosen to improve the ability to compare objective function values of signals of different lengths and contents.

5.3 Analysis of the optimisation problem

Prior to performing optimisation on data measured from a circuitc it is often of use to first analyse the optimisation problem to determine whether and where problems may arise. This step is divided into two areas: analysis using the objective function which is largely used in calibration, and analysis of parameter error which is used for parameter estimation.

5.3.1 Analysis using the objective function

The objective function is analysed for two purposes: the first as a method of confirming the results of the redundancy analysis, and the second to detect any parameters to which the function is insensitive.

By directly visualising the search space, it can be determined whether a global minimum exists, i.e. a single location at which ξ is at its minimum value, which for

simulated noiseless data will be $\xi = 0$, but will vary for real measurements. If one is found this indicates that the model has estimable parameters as no other combination of values could yield the minima.

'Screening' of parameters refers to the fixing of parameter values to reduce model complexity [100]. The technique is applied here to reduce the number of search space dimensions with the objective of reducing the optimisation time required to identify a circuit. Fixing a parameter that the objective function (and thus also the circuit model) is highly sensitive to would result in a high amount of error being introduced into the model, so to perform parameter screening first an analysis of the objective function is required to determine to which parameters it is least sensitive.

Objective function visualisation

Directly inspecting the shape of the objective function can provide immediate intuition into both the appearance of minima in the search space and the sensitivity of the objective function to each parameter. Figure 5.8 shows the top-down and side-on perspectives of the RC objective function using the simulated data set. Values of R_1 vary over $\pm 20\%$ of the specified accurate value, from which values of C_1 are chosen such that the product R_1C_1 remains the same.

Confirming the result of the redundancy analysis, the objective function of the RC circuit with no known components shown on the left of Figure 5.8 has a 'ravine', i.e. a curve along which $\xi = 0$. So long as the product R_1C_1 is the same as the accurate product, the transfer function and therefore objective function value is the same. The curve observed in the figure is due to the scaling: for R_1C_1 to remain the same when the value of R_1 doubles, the value of C_1 must be halved resulting in a curve instead of a linear relationship. Inspecting the objective function for sensitivity to the parameters, the objective function changes with similar magnitude with response to changes in R_1 and C_1.

The right side of Figure 5.8 is a plot of the objective function of the RC circuit with the addition of R_{ko}. Instead of there existing a curve of points over which $\xi = 0$ there is only one visible point, indicating that the transfer function of the accurate circuit cannot be produced by models with different parameter values. A second change to the shape of the objective function is that it is much more sensitive to changes in R_1 than to C_1, i.e. the value of the objective function changes faster with respect to changes in R_1 than C_1 as can be seen in the rapid change of colour in the y-axis but slow change in colour on the x-axis.

For models with more than 2 parameters it is not possible to plot the objective

function as a surface. Instead other visualisations must be utilised, for example a scatter plot of points with changes in the value of each parameter, projected onto an axis of just one parameter against the objective function value. This plot is demonstrated in Figure 5.9, where the objective function has been sampled using a grid of 7 points per parameter over a range of $\pm 20\%$ of R_1 and N, and values for C_1 and I_s chosen such that the products $R_1 C_1$ and $R_1 I_\mathrm{s}$ are maintained following the results of the redundancy analysis. The resultant 7^4 points have been projected onto an axis of R_1 to see how the objective function changes with respect to R_1. As with the RC circuit, both the SSDC with and without R_ko are illustrated.

For the SSDC model without R_ko there exist several minima, whereas with R_ko there exists only one. This may indicate that the redundancy in the original model is not present in the model with R_ko, but this indication cannot be assumed to be correct. Other minima may exist between the plotted points, and also in the other dimensions of the search space.

The second issue lies in the determining the sensitivity of the objective function to different parameters. Though, for the RC circuit, some intuition could be found on how much the value of the objective function changes with respect to changes in the parameter values, when only plotting the objective function against one parameter this is not possible. Multiple parameters must be plotted to then compare the sensitivity. Instead there exists a set of methods specifically designed for analysing a function's sensitivity.

Parameter Screening

Screening of model parameters requires a method of analysis which can be used to determine which parameters the objective function is most sensitive to, known as sensitivity analysis. Global sensitivity analysis (GSA) refers to the study of attributing uncertainty in a model's output (or this case the objective function) to uncertainty in a model's parameters and input. The prefix 'global' specifies that the analysis is upon the whole search space as opposed to local operating points. The implemented method, the Morris method [101], has been used to screen parameters in numerous fields, for example in the design of launch equipment for satellites [102]. *Trajectories* are generated that traverse the search space using a one-at-a-time strategy i.e. there is a change in only one parameter between neighbouring sample points. An *elementary effect* of a parameter can then be defined as

$$EE_i = \frac{\xi(\theta_1, ..., \theta_i + \Delta, ..., \theta_k) - \xi(\theta_1, ..., \theta_k)}{\Delta} \tag{5.39}$$

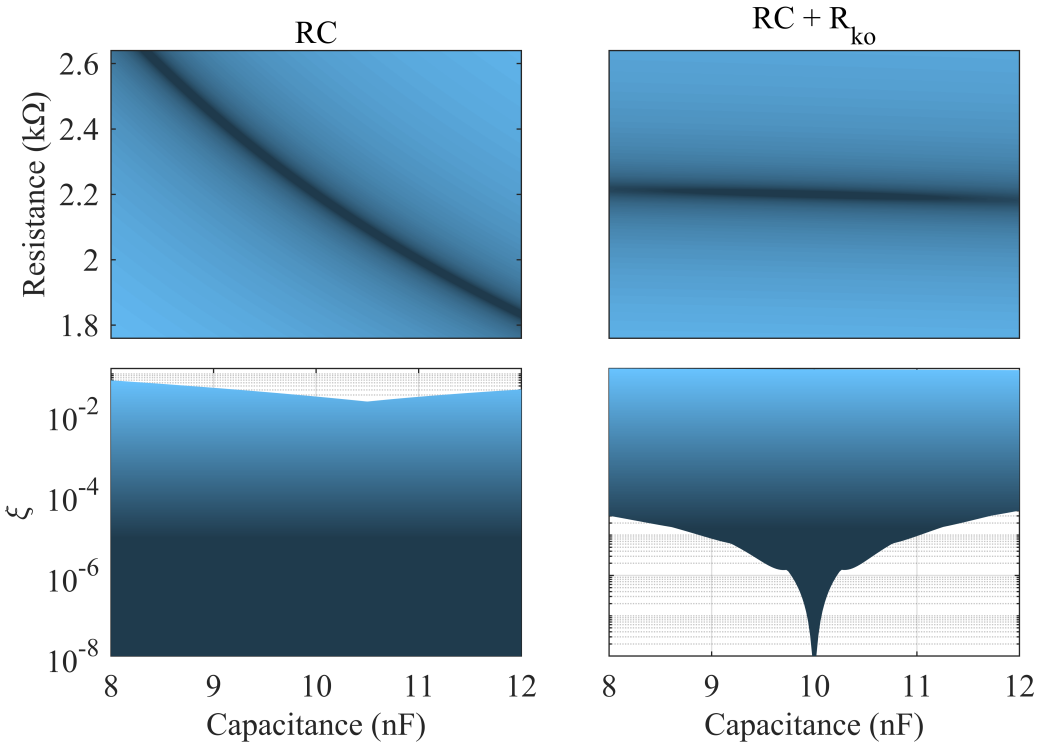


Figure 5.8: Top down and side-on views of the objective function surface of RC circuit, (left) RC model, (right) RC model with R_{ko} included. The base 10 logarithm of the objective function value is displayed to highlight the points where the value is zero (which has been substituted with 10^{-8} to avoid computation problems).

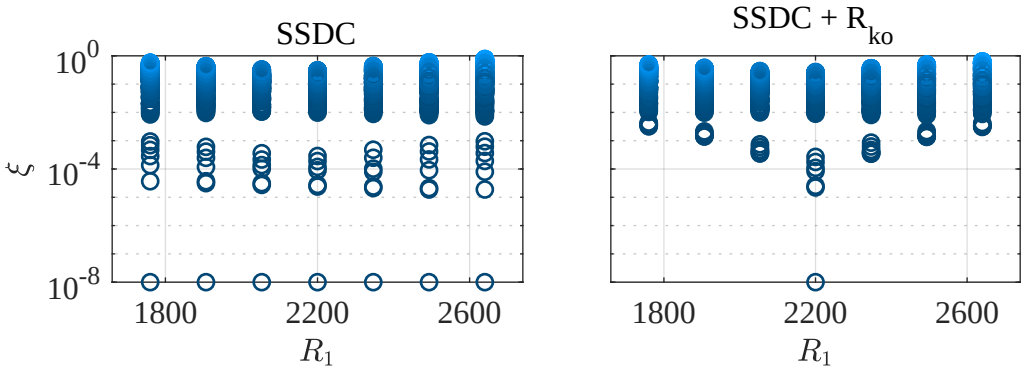


Figure 5.9: Scatter plot of the SSDC objective function. A range of 7 points for each parameter are chosen from $\pm 20\%$ of their value, except for C_1 and I_{s} which were chosen to maintain the products $R_1 C_1$ and $R_1 I_{\text{s}}$.

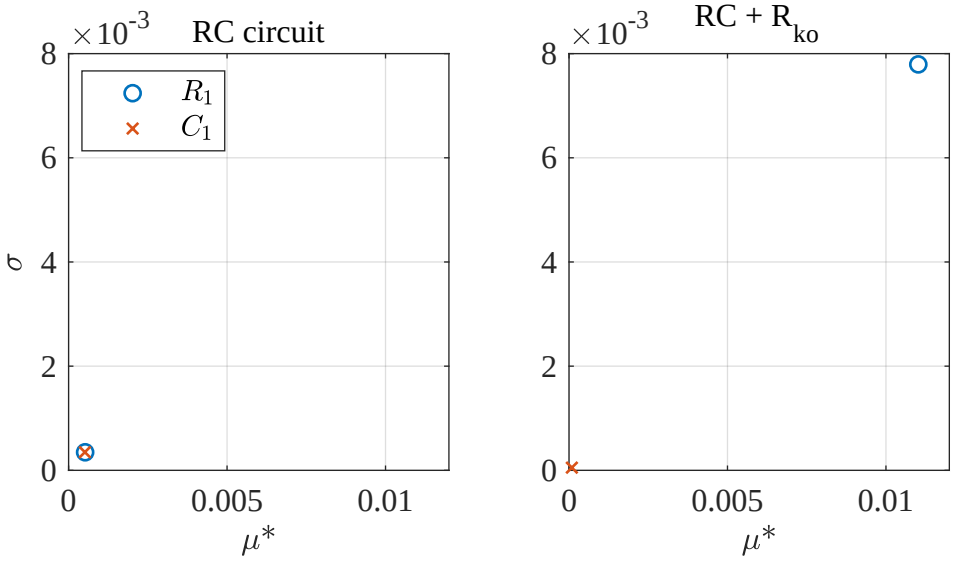


Figure 5.10: EE test on the RC circuit with and without R_{ko}, $r = 300$.

where θ_i is the parameter changed by the value Δ for the elementary effect. The number of calculated elementary effects for each parameter is given by the number of trajectories, r. The elementary effects are processed to create two sensitivity measures, μ^* and σ expressed by

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r}|EE_i^j|, \quad \sigma_i = \left(\frac{1}{r-1}\sum_{j=1}^{r}(EE_i^j - \mu_i)^2\right)^{\frac{1}{2}}. \tag{5.40}$$

The estimated absolute mean, μ^*, reflects the overall influence of the parameter on the objective function, and differs from the mean, μ, by using absolute values of the elementary effects, preventing type II errors which are caused by negative values [103]. The estimated standard deviation σ groups both the nonlinearity of the parameter and the dependence on other parameters relative to the change in the objective function. Intuitively, this can be understood by considering a change in the value of the elementary effects: the change must either be caused by a nonlinear parameter i.e. the effect changes across the range of parameter values, or by a change in another parameter due to sampling at other locations in the space.

GSA presented in this thesis was performed using SAFE, a MATLAB toolbox for Global Sensitivity Analysis [104]. Latin Hypercube Sampling was chosen from the toolbox, a method of sampling that divides the space into a grid from which near-random samples are chosen while avoiding repeated samples on a given hyperplane [105].

To confirm that the Morris method functions as expected with the circuit case stud-

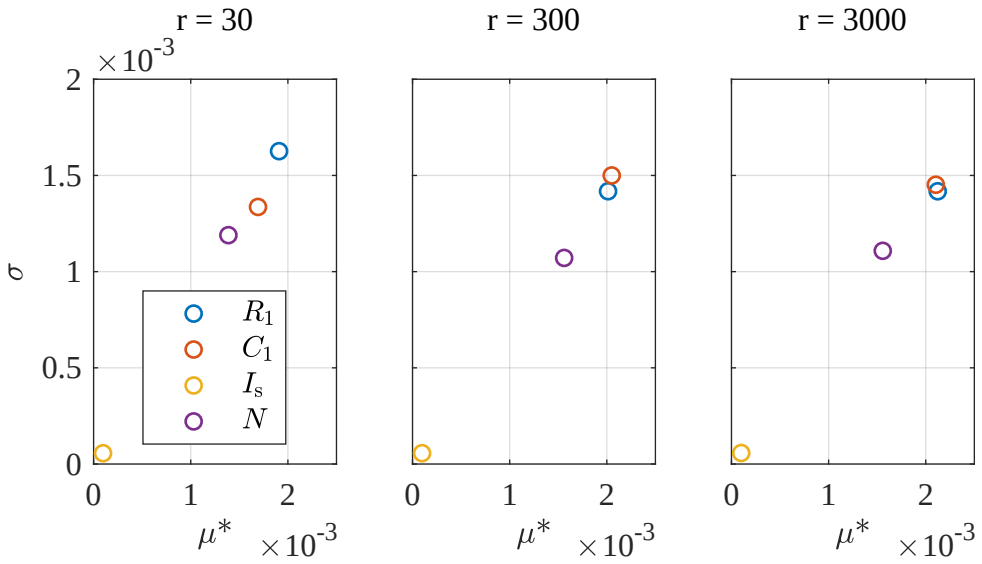


Figure 5.11: EE test on the SSDC with different numbers of trajectories, $r = 30,\ 300,\ 3000$.

ies, the RC circuit objective function can be analysed with and without R_{ko}. Figure 5.10 shows this comparison, with the left figure showing the results of the model without the known component, and the right with. As expected, the left plot shows the objective function to be approximately equally sensitive to both parameters, whereas for the right plot the objective function is more sensitive to R_1 than C_1.

Having confirmed comparable results from GSA and direct visualisation, the following question when using GSA is how many trajectories will correctly assess the sensitivity of each parameter. While this factor is dependent upon the model being analysed, an estimate can be found through testing different numbers of trajectories on a single model. Figure 5.11 shows the EE test on the SSDC without R_{ko} for three different numbers of trajectories. Comparing the plot for $r = 30$ to $r = 300$, there is a noticeable difference between the sensitivity of R_1 and C_1. With $r = 300$, the difference is visibly less, with R_1 and C_1 close than for $r = 30$. Broad trends have been largely captured using $r = 30$, however, I_{s} has a low sensitivity in comparison to the other three parameters.

Table 5.1 contains the sensitivity indices of the EE test with $r = 300$. Using the values of μ^* to compare the sensitivity of the objective function to each parameter, I_{s} is over 15 times lower than the rest of the parameters. As the objective function is so insensitive to I_{s} it makes a good candidate to remove as a parameter and instead fix at one value when optimising the model. By fixing a parameter, one dimension of the search space has been removed which should ideally reduce the time the optimisa-

Table 5.1: Sensitivity indices of the EET test on the SSDC, $r = 300$.

$\boldsymbol{\theta}_{\mathrm{dc}}$	μ^*	σ
R_1	$2.25e-3$	$1.50e-3$
C_1	$2.32e-3$	$1.58e-3$
I_{s}	$0.13e-3$	$0.06e-3$
N	$1.59e-3$	$1.15e-3$

tion takes, and if the objective function is insensitive to that parameter then the error introduced by fixing a parameter should be small.

To test this a set of optimisations were performed on 3 models of the SSDC, two featuring all parameters, with and without R_{ko}, and one with the value of I_{s} fixed at an initial random value, also without R_{ko}. The optimisation algorithm selected was the Nelder-Mead algorithm as implemented in MATLAB's `fminsearch` [95] with a maximum of 10000 iterations and an exit condition of a change in both function and parameter value of less than 1×10^{-25}. The results of this experiment are displayed in Figure 5.12, comparing the final value of ξ against the measured `cputime` from MATLAB, forcing the optimisation to use only one thread to ensure that the measured time is not skewed by multithreading.

Comparing the models without the known component, a marginal increase in error can be observed along with a decrease in computation time. The minimum mean computation time is achieved by the model using all parameters with the inclusion of R_{ko}, though the minimum recorded time is when I_{s} is fixed.

Linking the objective function to parameter error

In the design of electronic circuits one must be aware of the tolerances of given component's parameters. Component parameter values will vary between instances of a circuit, causing changes in the i/o behaviour between each circuit. In the parameter estimation strategy presented in this work, it is assumed that if the component parameters of a circuit are matched exactly then the i/o behaviour should be accurately reproduced, not accounting for non-ideal factors such as noise or parasitic impedances.

Common resistor tolerances are $\pm 20\%$, $\pm 10\%$, $\pm 5\%$, and $\pm 1\%$, with precision resistors starting at $\pm 0.1\%$ [106]. Capacitors can be given to similar tolerances though precision capacitors are uncommon in audio applications. Tolerances for the parameters of modern BJTs are typically much wider, for example with the 2N3906, the range

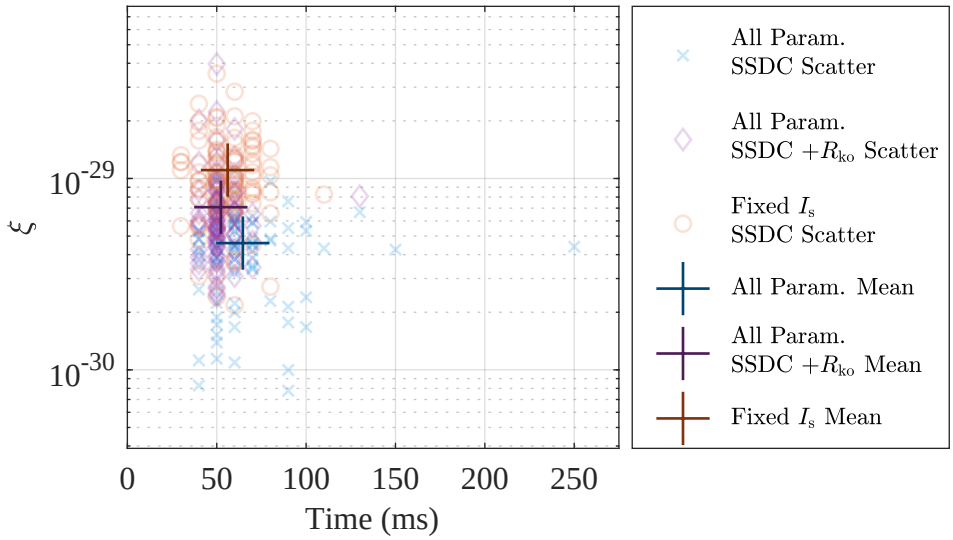


Figure 5.12: Comparison between optimising the SSDC with all parameters, with and without R_{ko}, and with I_{s} fixed. Computation time is compared against final ξ. 100 optimisations were performed with random parameter values selected from a uniform distribution over a range of $\pm 20\%$ of each accurate parameter value.

of DC current gain can take values between 100 and 300 [107]. For common diodes such as the 1N4148 or 1N4001, only maximum ratings are provided giving no value for total range of parameters [108, 109].

A survey of 3 guitar pedals show that resistor tolerances vary between $\pm 10\%$ depending on the value of the device[3]. Further, from the extracted parameters of the BJTs in Section 4.4.3, the worst case parameter tolerance is $\pm 70\%$.

This scenario can be used as a reference for the success of calibration. Figure 5.13 shows the relationship between $\langle \epsilon_\theta \rangle$ against ξ for the simulated SSDC, shown on a log-log graph to better highlight the data trend. Parameter error is defined in (5.41), and $\langle \cdot \rangle$ is used to note the mean of the error between each parameter in $\boldsymbol{\theta}$. A line of best fit has been found using MATLAB's `polyfit` using the log scale, producing the expression $\xi = 10^{1.99\langle \epsilon_\theta \rangle + 1.28}$. Computational limits are shown beneath $\langle \epsilon_\theta \rangle = 10^{-15}$, the resolution becoming insufficient to differentiate the results from zero, with these elements not shown as $\log(0) = -\infty$.

Three markers are placed across the plot indicating the mean component value error of 0.1%, 1%, and 10%. For the SSDC, 10% parameter error corresponds to $\xi \approx 0.24$, 1% to $\xi \approx 2.4 \times 10^{-3}$ and 0.1% to $\xi \approx 2.4 \times 10^{-5}$. A reduction in parameter error by

[3] `https://bholmesqub.github.io/thesis/chapters/identification-design/component-tolerances/`

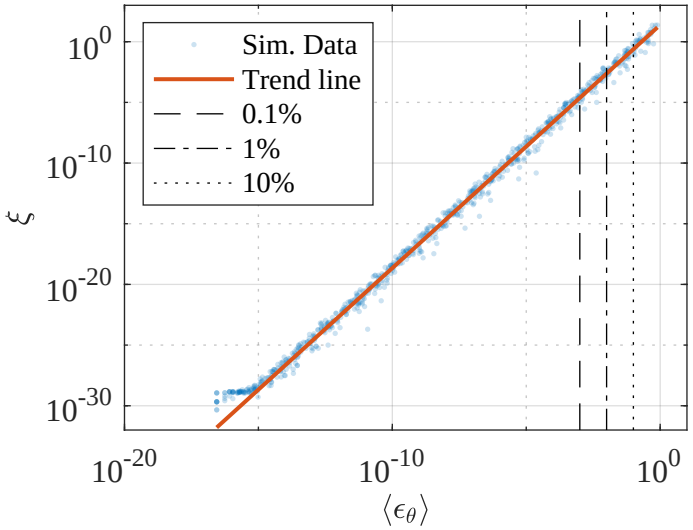


Figure 5.13: Objective function value ξ for the SSDC as a function of mean parameter error $\langle \epsilon_\theta \rangle$. The excitation signal used is that described in Section 5.2.1.

a factor of 10 is therefore relative to reducing the value of ξ by a factor of 100. Given that the ability of the available LCR meter used to directly measure components cannot measure to within 0.1% of a component's nominal value, the highest measurable change would be a decrease in ξ by 10^3, i.e. 10% to 0.1%. As nonlinear components can vary up to 100% in parameter value, a similar level of achievement would be a value of around $\langle \epsilon \rangle = 1\%$, or $\xi \approx 2.4 \times 10^{-3}$.

Validation signal comparison

To determine the improvement of a model's fit to the measurements of a circuit, validation metrics are required. In the optimisation procedure only one signal is used to capture all of the relevant data about the circuit, such that the simulation time for each iteration is minimal. When validating a model, the process is less time sensitive, and therefore a wider selection of signals can be used that better relate to the desired behaviour. As the objective function is designed such that the returned value is reflective of the error between signals independent of what that signal contains, different signals can be compared using the objective function to investigate the model's accuracy over different regions of operations.

The design of the excitation signal requires the specification of both a frequency and amplitude range. These ranges can be investigated using a set of new excitation signals for validation, each signal containing a single windowed sine wave with dif-

ferent peak amplitudes and frequencies, producing a map of objective function values that compare the measured response of the circuit against the model's response. The resultant map will illustrate where the greatest regions of error are with respect to peak amplitude and fundamental frequency.

Figure 5.14 contains an example comparison of maps for the SSDC. Four maps are displayed showing the objective function value surface comparing the simulated measurements against models with different parameter values, the first using the set used to seed the optimisation and the other 3 using parameters resulting from optimisations with different settings. A change of over 20 orders of magnitude is observed between the initial parameters and the optimised parameters, indicating that the optimisation has successfully improved the match between the simulated measurements and the model. These results using simulated data exceed the highest assumed accuracy as stated in Section 5.3.1. Differences between each of the optimised models are less marked, with the fit decreasing with frequency across each of the plots.

5.3.2 Analysis of parameter error

From the investigation of visualising the objective function in Section 5.3.1, several flaws were found in using the visualisation to demonstrate whether the parameters of a model are estimable. Instead of operating on the objective function value, the error in the estimated parameters can be inspected throughout the optimisation to determine whether they are converging to the accurate values, directly validating whether the parameters are estimable. A definition of parameter error is given by

$$\epsilon_\theta = \frac{|\hat{\theta} - \theta^*|}{\theta^*} \tag{5.41}$$

where $\hat{\theta}$ is the current estimate of the parameter, and θ^* is the accurate or reference parameter value. As the error is a magnitude normalised by θ^* the error will be 1 when $\hat{\theta} = 2\theta^*$ and $\hat{\theta} = 0$, and corresponds to a percentage of error if multiplied by 100. This metric requires a value with which to compare to, and as the method is designed to retrieve values without circuit disassembly this is not always available. To prove utility of the parameter estimation method, circuits will be disassembled to measure parameters directly thus providing a set of reference values.

When working with simulated measurements, the set of parameter values used in the model when generating data becomes the reference set. Testing the optimisation on simulated data prior to using measurements enables verification not only of whether the parameters are estimable, but also that the combination of objective function and

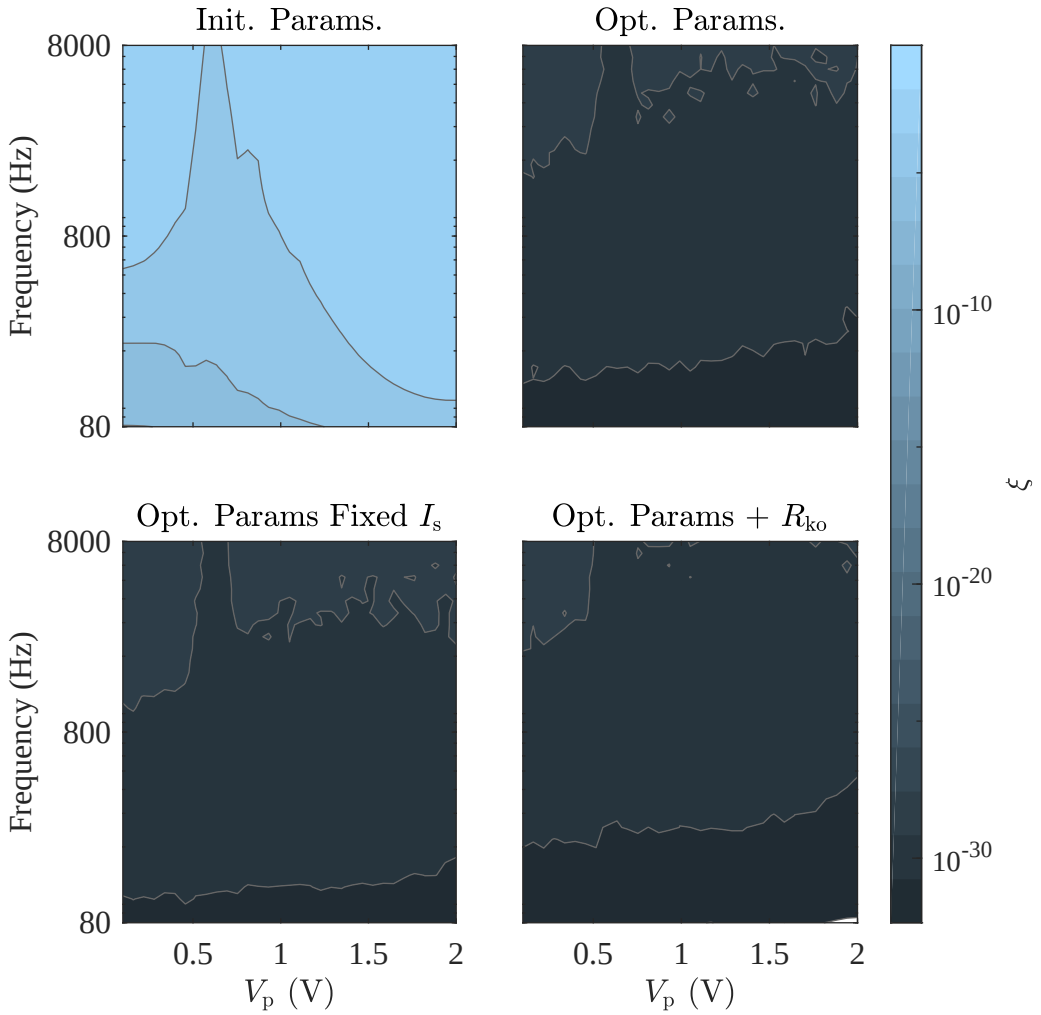


Figure 5.14: Four contour plots showing the SSDC objective function value for a set of Hann windowed sine waves of different peak amplitudes and frequencies, each plot using a different set of parameter values. (top left) The initial random parameters prior to optimisation, (top right) resulting values after optimising the model without R_{ko} and including I_s, (bottom left) resulting values after optimising the model with fixed I_s and without R_{ko}, (bottom right) resulting values after optimising the model with R_{ko} and including I_s.

excitation signal contain sufficient information to enable the parameters to converge, free of additional measurement challenges such as noise.

Figure 5.15 illustrates the convergence of the RC circuit model parameters for the models with and without R_{ko}. The optimisation was repeated 300 times of which the mean parameter error was taken across both optimisations and parameters, noted by $\langle \epsilon_\theta \rangle$, and is displayed logarithmically to better capture the data trend. Convergence is clearly observed for the RC model with R_{ko} in comparison to the model without it for which ϵ_θ is approximately constant, despite the fact that the maximum final objective function value across all optimisations was $6.82e-18$.

This again can be shown for the SSDC, with results shown in Figure 5.16. Similar results are displayed with the extended model achieving a much better convergence to the accurate parameter values than the model without R_{ko}. A significant difference is caused by the nonlinear diode model however, the parameter in the exponent N is always estimable as it cannot combine with other parameters, leading to some convergence being observed in the model without R_{ko} as seen in the left plot. Focusing on this parameter – as seen in the right plot – reveals that the parameter has a lower error than the average and follows the same trend as the model with R_{ko}, though does not achieve the same level of convergence.

To further understand the difference between circuits with and without R_{ko}, analysis can be performed on the 300 optimised sets of $\boldsymbol{\theta}_{\text{rc}}$, and $\boldsymbol{\theta}_{\text{dc}}$. For the RC circuit this analysis is demonstrated in Table 5.2, for the SSDC in Table 5.3. Alongside a comparison between the mean values of $\boldsymbol{\theta}_{\text{rc}}$ and $\boldsymbol{\theta}_{\text{dc}}$ and their error in percentage from the specified values, a new metric Relative Standard Deviation (RSD) has been included. RSD normalises the standard deviation of the set of optimised parameter values by the accurate value, removing the units from the result so each parameter can be compared, and is displayed as a percentage of the accurate value. For both circuits, the error of the mean value drops dramatically when including R_{ko}, with a correlated drop in RSD indicating that each optimisation converges to a small area of values.

Modelling modifications to circuits

A core reason to retrieve the parameter values of a circuit are that if the circuit is changed, the component parameters will be maintained independent of changes in topology or addition/removal of other components. One way of determining how successful a given set of parameter values may adapt to changes in a circuit is to change the load of the given circuit. Figure 5.17 shows this test for the SSDC with several different sets of parameter values. Let accurate parameter values be given by R_1^*, C_1^*,

Table 5.2: Optimised parameters of the RC, from left to right: the specified nominal parameter value, mean optimised value, error of mean value, and RSD of full set of 300 optimised values.

$\boldsymbol{\theta}_{\mathrm{rc}}$	Spec.	Mean	Error	RSD
Optimised RC				
R_1 (kΩ)	2.2	2.250	2.29%	12.1%
C_1 (nF)	10	9.913	$8.69 \times 10^{-1}\%$	11.7%
Optimised RC $+R_{\mathrm{ko}}$				
R_1 (kΩ)	2.2	2.200	$3.78 \times 10^{-12}\%$	$1.46 \times 10^{-10}\%$
C_1 (nF)	10	10.00	$6.75 \times 10^{-11}\%$	$3.24 \times 10^{-9}\%$

Table 5.3: Optimised parameters of the SSDC, from left to right: the specified nominal parameter value, mean optimised value, error of mean value, and RSD of full set of 300 optimised values.

$\boldsymbol{\theta}_{\mathrm{dc}}$	Spec.	Mean	Error	RSD
Optimised SSDC				
R_1 (kΩ)	2.2	2.191	$4.29 \times 10^{-1}\%$	13.7%
C_1 (nF)	10	10.30	3.02%	22.3%
I_{s} (fF)	10	1.028	2.81%	21.9%
N	1	9.999	$6.16 \times 10^{-3}\%$	10.7%
Optimised SSDC $+R_{\mathrm{ko}}$				
R_1 (kΩ)	2.2	2.200	$1.86 \times 10^{-13}\%$	$6.05 \times 10^{-13}\%$
C_1 (nF)	10	10.00	$7.11 \times 10^{-13}\%$	$6.26 \times 10^{-13}\%$
I_{s} (fF)	10	10.00	$5.29 \times 10^{-12}\%$	$2.22 \times 10^{-13}\%$
N	1	1.000	$2.11 \times 10^{-13}\%$	$9.42 \times 10^{-13}\%$

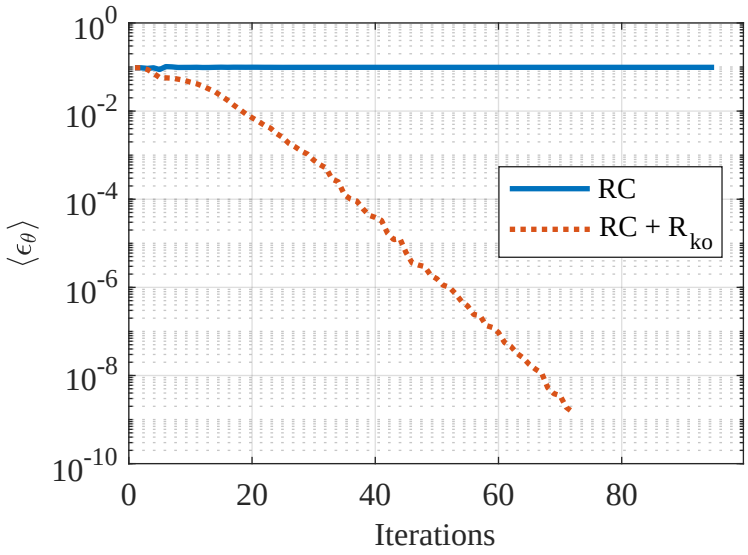


Figure 5.15: Mean parameter error against optimisation iterations for both the RC circuit model with and without R_{ko}. 300 optimisations were performed to produce the mean data. The maximum final objective function value between both models and all 300 optimisations was 6.82×10^{-18}.

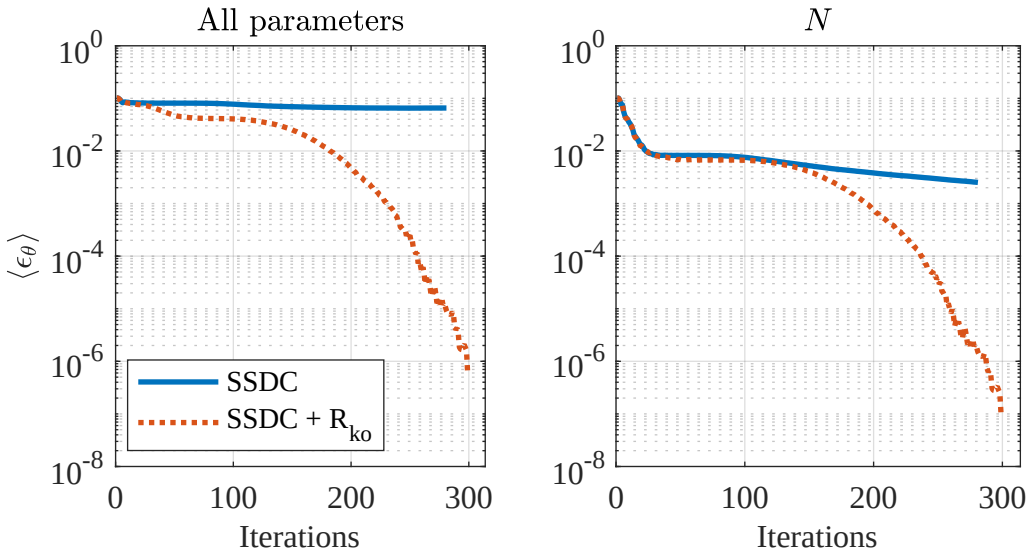


Figure 5.16: Mean parameter error against optimisation iterations for SSDC circuit model with and without R_{ko}. 300 optimisations were performed to produce the mean data. The maximum final objective function value between both models and all 300 optimisations was 5.38×10^{-6}.

I_s^*, and N^*, and their matching estimates be referred to as $\hat{R}_1$, $\hat{C}_1$, $\hat{I}_s$ and $\hat{N}$. Following the results of the redundancy analysis in Section 5.1.3, for the SSDC without R_{ko} equivalent behaviour can be found from different parameter values by letting

$$\hat{C}_1 = \frac{R_1^* C_1^*}{\hat{R}_1}, \tag{5.42}$$

$$\hat{I}_s = \frac{R_1^* I_s^*}{\hat{R}_1}, \tag{5.43}$$

$$\hat{N} = N^*. \tag{5.44}$$

The value of $\hat{R}_1$ has then been changed resulting in parameter sets that produce equivalent behaviour when unloaded, i.e. $R_{ko} = \infty$.

To demonstrate how much error is introduced for different load resistances, Figure 5.17 sweeps the load resistance against the resultant value of ξ where different amounts of error have been introduced to the parameter set. The curves are bell-like, at high load resistances the circuit becomes equivalent to being unloaded, and therefore fits exactly with the estimated parameters. At low values of load resistance, the output voltage becomes small as most of it is across R_1, meaning that the difference between accurate and estimated parameters would be small regardless of the parameter error. In the middle the error peaks due to there being a significant mismatch in behaviour. Although the parameters completely reproduce the behaviour of the original simulated circuit, when the load is changed the behaviour no longer matches. As the parameter error increases the corresponding value of ξ becomes larger, and vice versa should the estimated parameters be perfectly accurate no error would be exhibited with a change in load.

5.4 Conclusion

This chapter has introduced a framework for identifying audio circuits exclusively using input/output data by optimising the physical parameters of a circuit model. The framework was applied to maximally simple case studies, the linear RC circuit and nonlinear SSDC.

Redundancy was found in models of both circuits though they used different modelling paradigms, indicating a possible common redundancy in models of circuits that only express the input/output behaviour of a circuit. As only two cases were presented, further investigation is required to see if this finding holds for other circuits and modelling paradigms.

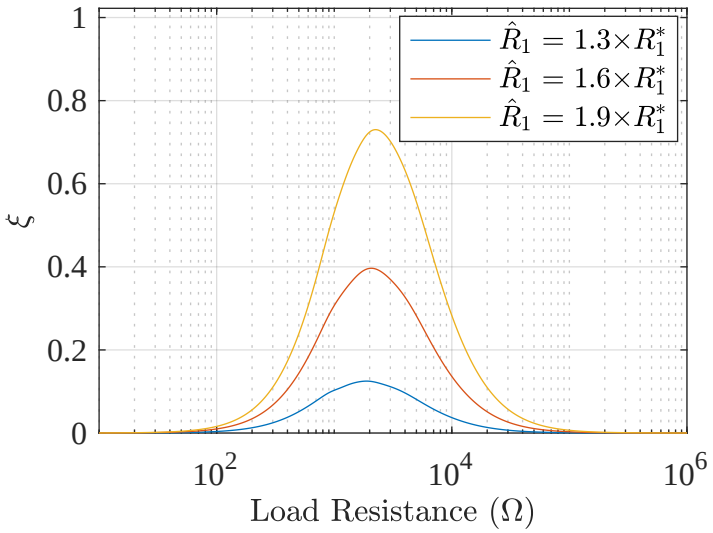


Figure 5.17: Objective function vs. different load resistances.

The detected redundancy was avoided by introducing additional information into the model in the form of a component with known parameters across the output terminals, in this case a resistor R_{ko}. Confirmation of this redundancy was shown through a variety of direct visualisation of the objective functions and plotting the convergence of the parameters during the optimisation process, though primarily the latter. Optimising the RC circuit without R_{ko} showed no convergence in parameter value, whereas with R_{ko} a high level of accuracy was achieved. With the SSDC, the result was less dramatic as the parameter N was always estimable, though a significant improvement in final parameter value accuracy was noted when using R_{ko}.

Though the initial derivations of the circuit models exhibited parameter redundancy, validation of model's with and without R_{ko} showed comparable behaviour across the desired operating range when compared using the sets of sine waves of varying amplitudes and frequencies. Further, through performing the EET on the SSDC it was found that one of the parameters could be fixed at an incorrect value and still yield a similar final value of ξ. A minor decrease in computation time was observed, and although the optimised parameters for the model using R_{ko} was on average faster to compute, this result could indicate that on models with a higher number of parameters that more parameters could be fixed which may yield more of a computational saving. Further value could be derived from this strategy for models with such a high number of parameters that the optimisation is prohibitively expensive.

Chapter 6

Calibration of the Dallas Rangemaster model

Calibration has been defined within the scope of this thesis as tuning of a physical model to best match the measured behaviour of a circuit over a given range of operation. To achieve this match, a minimum of user time and effort is desirable: minimal time spent performing measurements, designing models, and waiting for the optimisation to produce the calibrated model. Minimising time and effort facilitates the application of calibration to new circuits, removing barriers from immediate results.

In this chapter, the Dallas Rangemaster Treble Booster guitar pedal is calibrated. The primary objective is to determine with what accuracy can the circuit be identified given the restrictions defined by calibration. Identification is here performed on measurements of a real instance of the circuit, as opposed to only simulated measurements demonstrated throughout Chapter 5.

There are two secondary objectives around the calibration process. The first relates to the BJT analysis from Chapter 4 to further investigate the differences of semiconductor material. As the original Dallas Rangemaster used a germanium OC44, a second BJT, the silicon BC557, is also used in the circuit to provide a point of comparison. By comparing the results of the identification, the objective is to further investigate how successfully each BJT can be modelled within the context of the circuit using the Ebers-Moll model.

Additionally the EET introduced in the previous chapter is again applied to screen model parameters. As the Rangemaster uses more parameters than the single sided diode clipper, the circuit model provides potentially a better case for investigating the effects of fixing parameters to improve optimisation speed.

This study is adapted from the DAFx 16 publication [110], but updated using the

Table 6.1: Parameters used in the calibration of the Dallas Rangemaster. Nominal values are taken from the schematic, measurements, and SPICE models.

Parameter	Units	Value		
		Nominal	Measured	Stochastic Sample
R_1	kΩ	470	473.3	506.6
R_2	kΩ	68	68.60	63.03
R_3	kΩ	3.9	3.897	4.602
R_4	kΩ	10	10.00	9.755
R_o	MΩ	1	0.997	0.813
C_1	nF	4.7	4.92	4.477
C_2	µF	47	46.95	51.99
C_3	nF	10	11.57	11.18
OC44				
I_s	µA	3	$2.7 - 3.2$	2.249
N	–	1	-	0.991
β_f	A/A	100	$98.4 - 214.8$	95.65
β_r	A/A	10	$9.6 - 24.5$	11.17
BC557				
I_s	fA	40	-	-
N	–	1	-	-
β_f	A/A	340	-	-
β_r	A/A	15	-	-

results of the research into germanium BJTs from [97] and successive research carried out by the author into circuit identification. Corrections have been applied to the initial parameter values for germanium BJTs. The EET has been re-purposed to screen parameters instead of ranking their sensitivity. Discussion of parameter error has been removed to improve the focus on calibration.

6.1 Identification design

6.1.1 Model and parameter selection

As has been used throughout the thesis, a state-space model was selected to represent the Dallas Rangemaster for the identification, it's topology defined by the schematic in Figure 3.4. Maximally simple component models were selected to define the initial parameter set; single-parameter linear component models and the Ebers-Moll model from (2.5, 2.6) form

$$\boldsymbol{\theta}_{\mathrm{rm}} = [R_1,\ R_2,\ R_3,\ R_4,\ R_{\mathrm{o}},\ C_1,\ C_2,\ C_3,\ I_{\mathrm{s}},\ N,\ \beta_{\mathrm{f}},\ \beta_{\mathrm{r}}]. \tag{6.1}$$

Nominal and measured parameter values for the model are displayed in Table 6.1. Nominal values for linear component parameters were taken from Table 3.2 (as have been used throughout) and were used to select the values for the physical circuit components. Values for the OC44 were informed by the extracted Ebers-Moll parameters from Chapter 4, and values for the BC557 informed by those used in LTspice. Measured linear component parameter values were taken from a multimeter with each component measured in isolation. The final column of the table contains a sample of randomly generated parameters taken from a uniform distribution where linear component parameters varied $\pm 20\%$ and BJT parameters varied $\pm 40\%$ of their respective nominal values. This sample is used in the generation of simulated input/output data that is later used in the validation of the identification design.

6.1.2 Excitation signal and objective function

The chosen range of behaviour for the model to capture is defined using the excitation signal. For the Dallas Rangemaster amplitude and frequency ranges were chosen to resemble the content of a guitar-type signal, as this is the anticipated use-case. The largest computational expense in optimising the behaviour of a model is the simulation of that model. Therefore, to minimise the computation time, a short excitation signal should be used.

To create a guitar-like excitation signal of minimal duration, signal parameters were chosen: $f_{\mathrm{s}} = 100\,\mathrm{kHz}$ and $N_{\mathrm{s}} = 20 \times 10^3$ producing a signal 0.2s long. Setting $d_{\mathrm{l}} = 10$ and $d_{\mathrm{h}} = 400$ selects a frequency range of $50\,\mathrm{Hz}$ to $2\,\mathrm{kHz}$, combined with $V_{\mathrm{p}} = 2\,\mathrm{V}$ to cover the anticipated guitar signal properties. To maximise the SNR of the output the inverse transfer function of the linearised Rangemaster was applied to the excitation signal. The resulting excitation signal and measured output signals

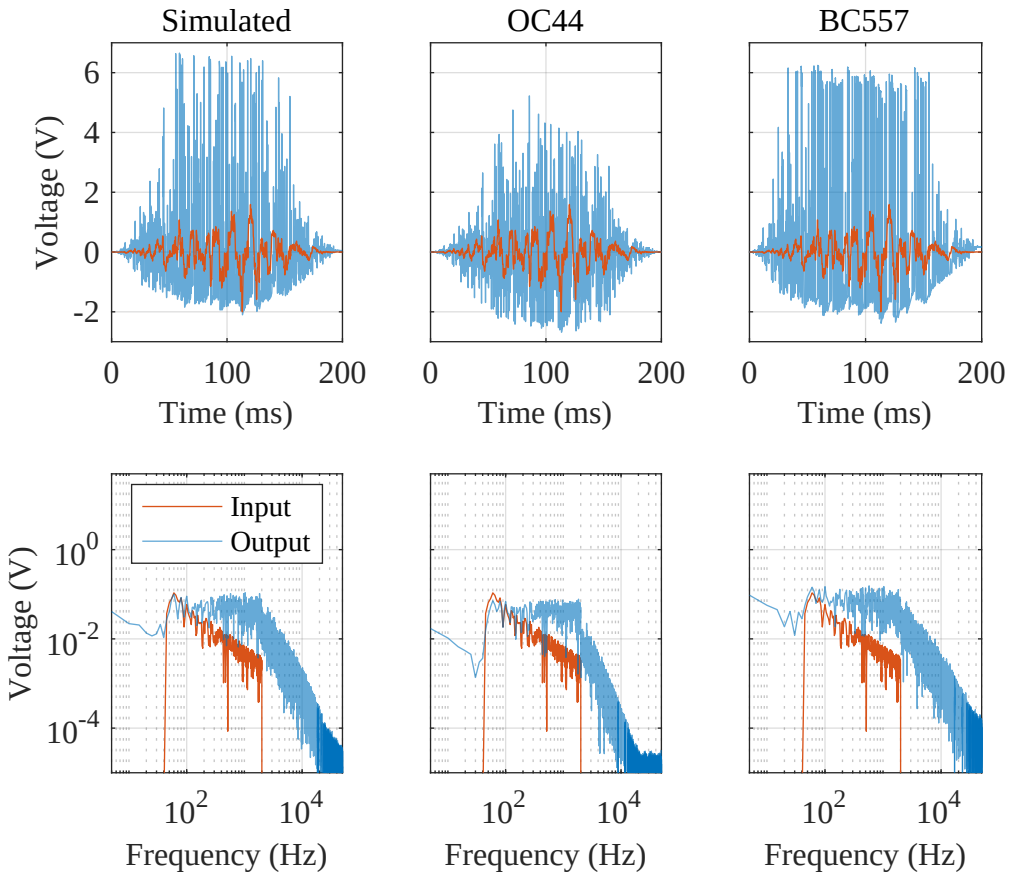


Figure 6.1: Input and output signals of the Dallas Rangemaster circuit for simulated measurement, OC44 circuit, and BC557 circuit.

from each of the OC44 and BC557 circuits and also the simulated measurement output signal using the stochastic parameters are illustrated in Figure 6.1.

At the time of the research the values of A_d used in the multi-sine excitation signal were not modified to ensure that their sum is equal to 1. Instead they were normalised using $N_\mathrm{s}/\sqrt{d_\mathrm{h} - d_\mathrm{l}}$, which provided a signal with a standard deviation that is constant and independent of the signal parameters. If $\sum_{d=d_\mathrm{l}}^{d_\mathrm{h}} A_d \neq 1$, then there are scenarios in which a high crest-factor can arise, but for the excitation signal used in this work – prior to the application of the Hann window – the crest factor was 3.36, indicating a low peak-to-peak voltage.

To compare the output of circuit and model the same objective function from (5.37) was applied.

6.1.3 Measurement setup

The Dallas Rangemaster circuit was assembled on a breadboard to facilitate the changing of BJTs and also the direct measurement of each linear component value. Identical circuits were used for the OC44 and BC557 BJTs, where the BJT was simply exchanged between measurements. The breadboard in question was a component of a National Instruments ELVIS II Data Acquisition system (DAQ) which offered 16 bit analogue inputs and outputs to drive and measure the circuit, and also a constant powersupply to drive V_{c}.

The input/output signals used in the identification process were sampled at $f_{\mathrm{s}} = 100\,\mathrm{kHz}$ and the output was averaged 100 times to reduce noise. Further signals were measured for the validation (discussed later in the chapter) and were measured at $f_{\mathrm{s}} = 400\,\mathrm{kHz}$ as the simulation time of the model is less critical during the validation process.

6.1.4 Optimisation algorithm

The majority of the optimisation was performed using the Nelder-Mead simplex method as implemented in MATLAB's `fminsearch` function and described in [95]. From an initial set of optimisation to test exit conditions, convergence was specified by a change smaller than 1×10^{-6} in both objective function and parameter values.

To further remove user-interaction from the process, a genetic algorithm (GA) was implemented to select starting points for the Nelder-Mead method. This was implemented using the `ga` function from MATLAB's global optimisation toolbox (for a thorough background on the method see e.g. [111]). Each iteration of the algorithm samples 100 parameter sets, beginning with values taken from a uniform distribution of parameter values, the limits being $\pm 20\%$ of the nominal linear parameter values and $\pm 40\%$ of the nominal BJT parameter values. The objective function value is found for each set, after which the parameter value set with the lowest corresponding objective function value is used as the starting point of the Nelder-Mead method. Constraints were placed on the GA solver to keep parameter values between the limits used for the original parameter sets, though the Nelder-Mead method was not constrained.

Successive iterations were formed from a combination of *crossover* and *mutation*. Crossover sets are generated by taking two sets from the previous iteration and randomly selecting parameter values from each set. Conversely sets generated using mutation take a single set from the previous iteration and change each value stochastically.

The 5 parameter sets with the lowest objective value roll over to the following

iteration without alteration. Of the remaining sets, 70% are taken from the crossover sets, and 30% are taken from mutation. As the sets which have been optimised using the Nelder-Mead method are likely to be the lowest for following iterations, these sets are prevented from repeatedly seeding the Nelder-Mead algorithm. Should the result of the Nelder-Mead method not be lower than the previous result for 5 iterations, the optimisation process is terminated assuming that the probability of finding a lower set is small.

6.2 Analysis of the optimisation problem

6.2.1 Objective function vs. parameter error

As in Section 5.3.1, here the objective function value ξ is linked to error in the parameter values, given by the mean error $\langle \epsilon_\theta \rangle$. This can be used to quantify how successful the calibration is: without concern to the actual component values, the resultant model can be compared to manufacturing an equivalent circuit with values within a given tolerance, as indicated by the final value of ξ.

Shown in Figure 6.2 are the results of introducing parameter error to component parameter values for the simulated model, and inspecting the corresponding values of ξ. The excitation signal is that described in Section 6.1.2. The line of best fit, or trend line, can be expressed as $\xi = 10^{2.01\langle \epsilon_\theta \rangle + 1.24}$, highly similar to that found for the SSDC in Section 5.3.1. Reducing the value of ξ by a factor of 100 reduces parameter error by a factor of 10.

As there is the nonlinear BJT with parameters that can vary over 70% of the mean measured values[1], the objective would be to reduced the effective parameter error to 1%, or $\xi \approx 3.4 \times 10^{-4}$.

6.2.2 Validation on simulated data

Prior to attempting to model a real circuit from measurements it is useful to first identify a system that is the model itself, i.e. generate simulated measurements from the model to use in the identification. Simulated measurements are free from noise and are guaranteed to be able to be fit by the model, providing an extreme best-case scenario for what the results of the optimisation on the real circuit may be.

[1] `https://bholmesqub.github.io/thesis/chapters/identification-design/component-tolerances/`

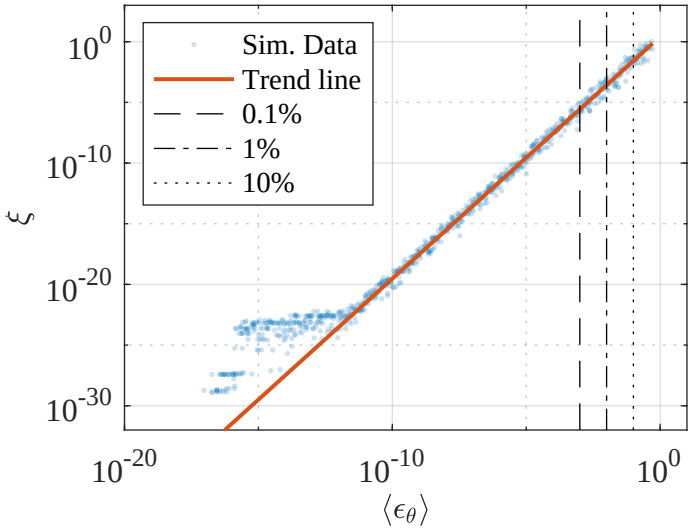


Figure 6.2: Objective function value ξ for the Rangemaster as a function of mean parameter error $\langle \epsilon_\theta \rangle$. The excitation signal used is that described in Section 6.1.2.

A single optimisation process was run using the Nelder-Mead algorithm using measurements generated with the stochastically sampled parameters from Table 6.1. Initial parameter values were taken from a uniform distribution with identical ranges as used for the parameters values used in the model that generated the data.

Figure 6.3 shows a contour plot consisting of the objective function values of a set of windowed sinusoidal signals with different amplitudes and frequencies. The amplitude and frequency ranges have been extended over that used with the excitation signal to ensure that the model still performs adequately in regions not explicitly covered in the specified ranges. The new ranges are $20 - 3000\,\mathrm{Hz}$ and $0.1 - 3\,\mathrm{V}$, each signal still windowed with a Hann window and with a duration of $200\,\mathrm{ms}$.

The left-hand plot shows the fit between the simulated measurements and the model using the initial parameter values. The right-hand plot instead uses the resulting parameter values from the Nelder-Mead optimisation. Values of ξ are approximately 8 orders of magnitude smaller for the optimised parameters than the initial parameters, indicating a successful capture of the simulated behaviour across the desired range of operation.

6.2.3 Parameter screening

The number of dimensions in the search space that is traversed by the optimisation algorithm is determined by the number of parameters in a model. In the previous chap-

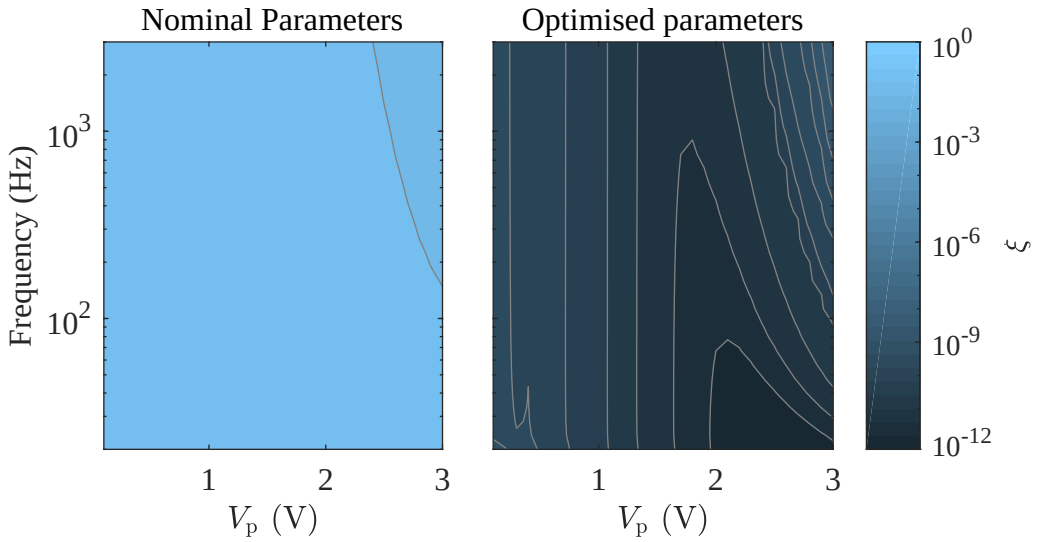


Figure 6.3: Validation of the Dallas Rangemaster optimisation using simulated data. The displayed surface represents the value of ξ for windowed sine waves of different amplitude and frequency, the left using a model with random initial parameter values, and the right using optimised values.

ter it was demonstrated that reducing the search space by one dimension a reduction in computation time could be achieved with a trade-off for the optimised objective function value. This approach is applied to the Dallas Rangemaster, using the EET to screen the 12 parameters.

Figure 6.4 contains the EET of the OC44 Rangemaster using the same method as described in Chapter 5 with $r = 300$. Parameters display a wide range of sensitivity indices, with several appearing relatively insensitive.

To compare the effects of fixing parameter values, three boxes have been placed around clear groups of parameters that are visually clustered, from which 4 different combinations are formed from fixed and variable parameters. A control 'combination' is included which uses all of the parameters. Two of the combinations fix the least sensitive parameters and leave only the parameters which the objective function is most sensitive to: the top 7, and the top 3. Finally to provide some context as to the effect of the number of parameters optimised and the sensitivity of the objective function to those parameters, the middle 4 parameters are chosen as the last combination. Each of these groups are to be compared in Section 6.3.2 for the computation time and resulting values of ξ of multiple optimisations using the Nelder-Mead algorithm.

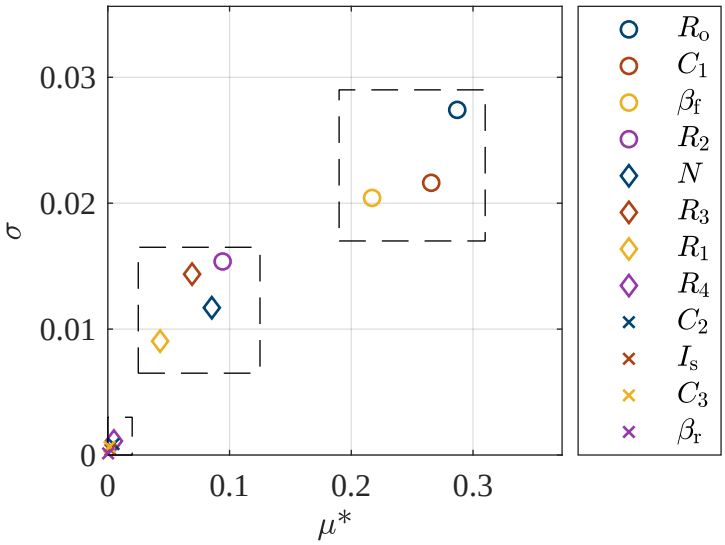


Figure 6.4: EET performed on the Dallas Rangemaster OC44 objective function, $r = 300$. Parameters have been boxed to indicate the selected groupings for screening.

6.3 Results and validation

Using the measurements taken from the Dallas Rangemaster, identification of the circuits was performed using the GA solver and Nelder-Mead algorithm. The resulting objective function and parameter values from the optimisations are displayed in Table 6.2. The OC44 Rangemaster is the primary objective for the identification. Referring to Figure 6.2, the resulting values of ξ correspond to $\langle\epsilon\rangle \approx 2.3\%$ for the germanium and $\langle\epsilon\rangle \approx 4.2\%$ for the silicon. While these haven't achieved the previously stated objective of 1%, the reduction in error is still satisfactory considering the initial parameters began over a range of $\pm 20\%$ for linear components and $\pm 40\%$ for the BJT parameters.

To expand upon the fit to the model validation results are displayed in Figure 6.5. The left plot shows the fit of the model with initial parameter values and the right shows the equivalent with the optimised parameter values. Clearly the change in fit is markedly increased across the desired range of operation.

Though an increase in fit to circuit behaviour has been observed, comparing the final parameter values in Table 6.2 to their measured equivalent in Table 6.1 it is clear the parameters have not converged. Accurate parameter values are not the primary objective, but this result does suggest parameter redundancy which could indicate the suitability of fixing parameters to reduce computation time. A parameter that the objective function is notably insensitive to is C_2, which has decreased in value by over 20

Table 6.2: Final objective function and parameter values for the OC44 and BC557 Rangemaster circuits.

	OC44	BC557
Final ξ	1.476e-3	4.587e-3
R_1	$2.061\,\mathrm{M\Omega}$	$107.9\,\mathrm{k\Omega}$
R_2	$52.45\,\mathrm{k\Omega}$	$43.52\,\mathrm{k\Omega}$
R_3	$635.1\,\Omega$	$9.571\,\mathrm{k\Omega}$
R_4	$96.75\,\mathrm{k\Omega}$	$3.314\,\mathrm{M\Omega}$
R_o	$35.87\,\mathrm{k\Omega}$	$10.56\,\mathrm{k\Omega}$
C_1	$5.842\,\mathrm{nF}$	$9.887\,\mathrm{nF}$
C_2	$62.52 \times 10^{-27}\,\mathrm{F}$	$35.74\,\mathrm{\mu F}$
C_3	$78.51\,\mathrm{nF}$	$4.111\,\mathrm{nF}$
I_s	$4.284\,\mathrm{\mu A}$	$318.8 \times 10^{-21}\,\mathrm{A}$
N	0.8434	1.432
β_f	15.41	174.5
β_r	1.264	113.3

orders of magnitude. This finding correlates with the EET that placed C_2 in the group of parameters of which the objective function is least sensitive.

Sound examples for the Rangemaster models and circuit can be found online[2] for readers to further asses the success of the calibration.

6.3.1 OC44 vs BC557 comparison

Comparing the final values of ξ between OC44 and BC557 Rangemaster circuits, the fit of the germanium OC44 is approximately 3 times better than the silicon BC557. Validation maps for the BC557 are displayed in Figure 6.6. The minimum of the OC44 map is at a lower value than the minimum of the BC557 map as might be expected given their respective optimised values of ξ. Further, the overall fit of the OC44 map is better over much of the range of operation.

To further investigate the differences in fit, it is useful to directly inspect a waveform. Figure 6.7 shows a sample of Rangemaster output when driven by the excitation signal, for the circuit and model of both the OC44 and BC557. The middle plots report the error between measured and modelled signals and lower plots the spectrogram of

[2]https://bholmesqub.github.io/thesis/chapters/calibration/

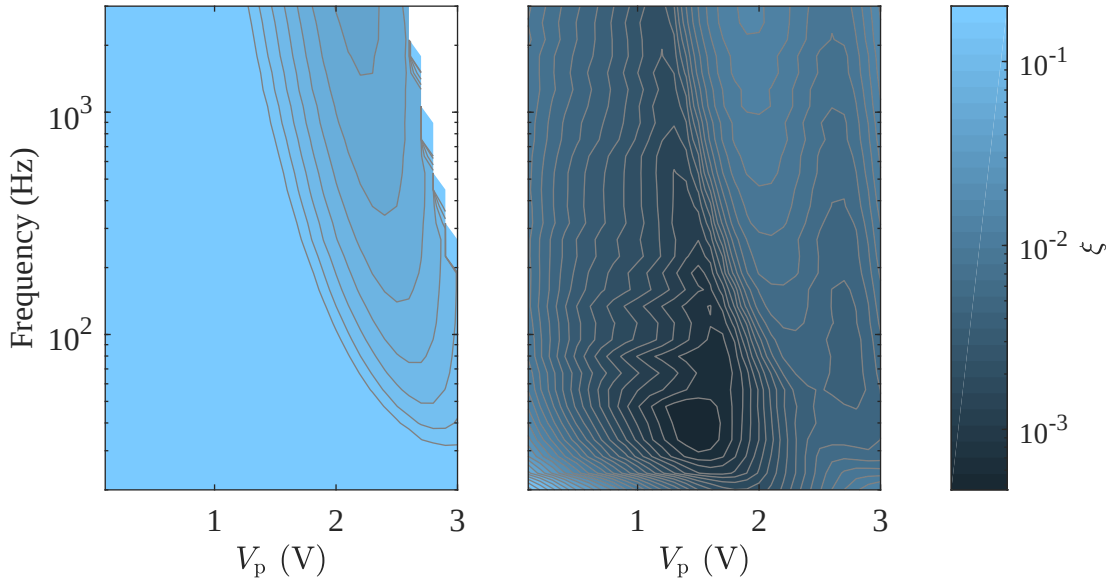


Figure 6.5: Error between measured and modelled output of the **OC44** Rangemaster in response to windowed sine waves at different amplitudes and frequencies. (left) Model uses nominal parameter values, (right) model uses optimised parameter values.

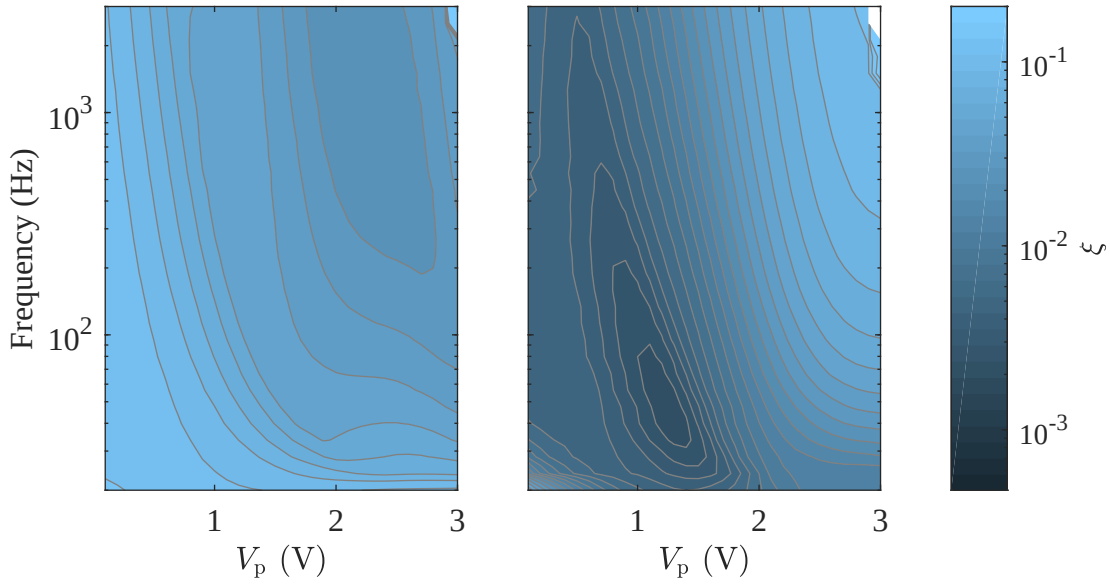


Figure 6.6: Error between measured and modelled output of the **BC557** Rangemaster in response to windowed sine waves at different amplitudes and frequencies. (left) Model uses nominal parameter values, (right) model uses optimised parameter values.

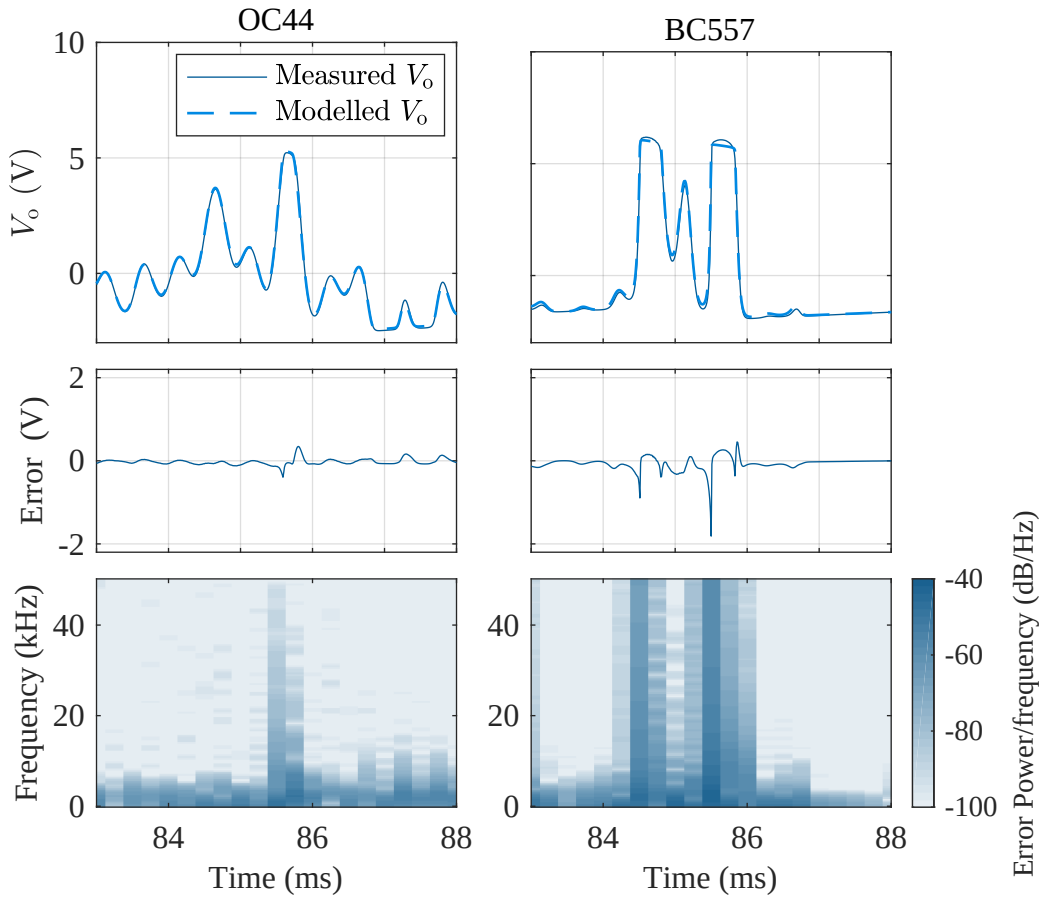


Figure 6.7: Comparison of a high-error section of the output of the Rangemaster when driven with the excitation signal, for both OC44 and BC557 BJTs. (top) Output voltages, measured and modelled, (middle) error between measurement and model, (bottom) spectrogram of error.

this error, both of which again indicate a higher amount of error for the BC557 than the OC44. Though both circuits are driven by the same signal there is an obvious difference between outputs. The BC557 circuit exhibits a higher gain that is observable through the fast transitions between the saturation regions at the top and bottom of the signals. These fast transitions cause large spikes in the error which points to a failure to capture the high frequency behaviour of the circuit as suggested by the spectrogram in the lower plot. Results of Chapter 2.1.1 indicated that the capacitance across the BJT junctions causes a change in high frequency response in the Rangemaster which could be a cause to the mismatch in high frequency here.

6.3.2 Parameter screening

The results of 20 optimisations performed using 4 different groups of parameters are displayed in Figure 6.8. Computation time was measured using MATLAB's `cputime` function as used in the previous chapter. The mean final objective function value has increased only a small amount between optimising all of the parameters and only using the top 7, though the mean computation time has reduced by more than double. Further reducing the group to only the top 3 parameters significantly increases the mean final objective function value, with a similarly extreme reduction in computation time. The last group included in the experiment is the middle 4 parameters. This group is placed between the top 3 and top 7 for both iterations and final objective function value.

To validate the investigation into parameter screening, the sine wave maps of the models using the closest-to-mean set of parameters for each of the four groups is illustrated in Figure 6.9. Generally the sine wave maps correlate with the mean final values of ξ. As might be expected the surface generated using all parameters has the lowest point, though when using the top 7 parameters a better fit is achieved in the low amplitude region.

6.4 Conclusion

This chapter has presented a calibration of the Dallas Rangemaster Treble Booster guitar pedal circuit. The identification was successful in that a significant increase in fit between the measurements from the OC44 circuit and model can be observed when comparing the model from before to after calibration.

A comparison between a silicon and germanium Rangemaster was also presented between the germanium OC44 and silicon BC557 BJTs. The identification of the circuit using the BC557 was less successful than that of the circuit using the OC44. A hypothesis of Chapter 4 was that a silicon BJT would be easier to identify within a circuit as the behaviour is closer to the ideal behaviour as represented by the Ebers-Moll model. This has shown to not necessarily be true because the OC44 Rangemaster achieved a lower final value of ξ than the BC557 Rangemaster, as well as a general better fit across the range of operation as shown by the validation maps. A caveat for this result is that as the gain of the OC44 is lower than that of the BC557 less high frequency distortion is present in the excitation and validation signals which may make it easier for the model to fit the measurements. Should the BC557 Rangemaster have been driven at a lower amplitude it is possible a similar level of fit to the OC44

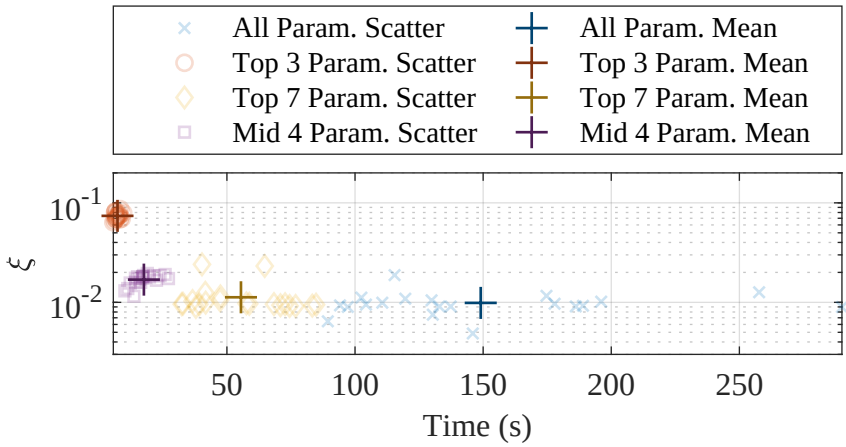


Figure 6.8: Iterations to complete optimisation vs final ξ value for the 20 optimisations of the germanium Rangemaster with different fixed parameters. Four combinations are presented using the groupings from the EET: all parameters, the top 7 and top 3 parameters, and the middle 4 parameters.

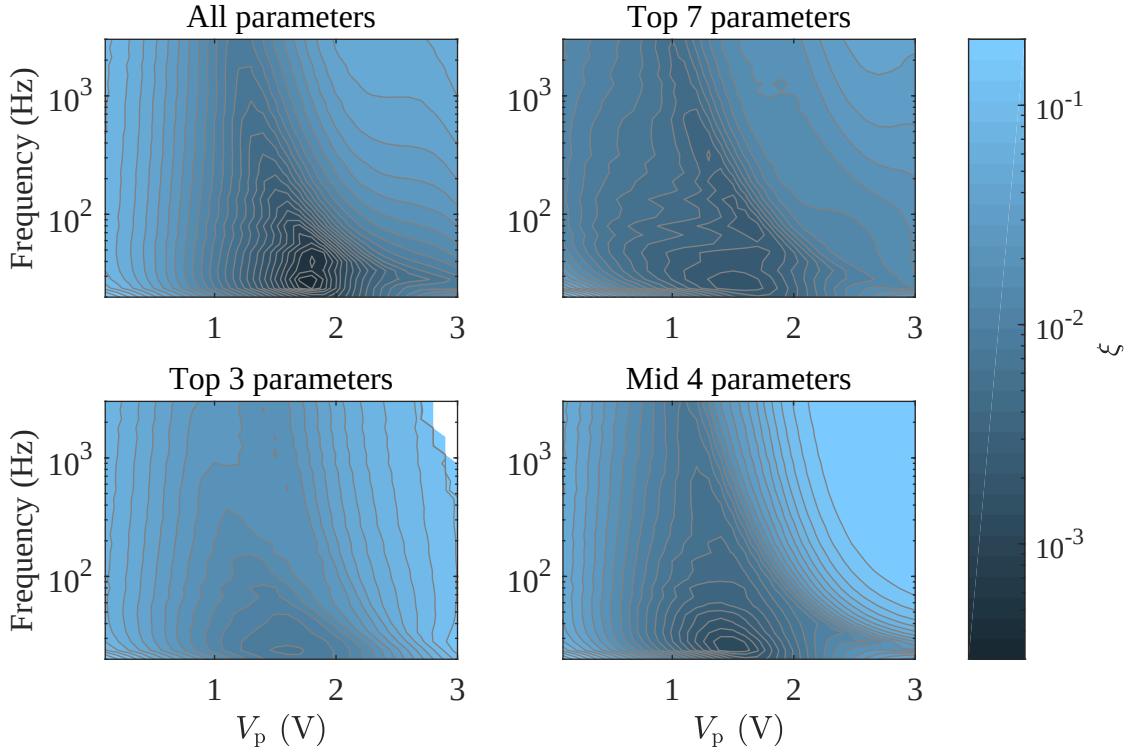


Figure 6.9: Comparison of models using optimised parameter values where different numbers of parameters were optimised. The displayed surface represents the value of ξ for windowed sine waves of different amplitude and frequency.

Rangemaster would have been achieved.

Additionally, as parameter values are not of interest in calibration, the potential non-ideal behaviour of the OC44 may have been compensated for by adapting the rest of the circuit's behaviour. This would mean that the BJT is still more difficult to identify within a circuit, but that the Rangemaster contains sufficient flexibility to compensate, and that it is easier to compensate for the OC44 than the BC557.

Through parameter screening an evident speed-up in the optimisation has been observed. By fixing 5 parameters, the computation was reduced on average by more than half, with a marginal increase in mean error. Comparing the validation maps for models optimised using the full set of parameters, and that using only the top 7 parameters, shows a comparable fit. However, the model optimised with all parameters shows a lower minimum error. This result suggests that screening parameters is a useful strategy to apply when calibrating circuit models, and recalling the marginal reduction in computation time of the single sided diode clipper from Section 5.3.1, further implies that when working with circuits with a higher number of parameters, it could have even more importance.

Comparing the middle 4 parameters to the top 3 parameters revealed that although the sensitivity of the objective function to a parameter may be important, in this case the number of parameters optimised is more significant. Both the final objective function values and validation maps revealed that the model using parameter values of the optimised middle 4 parameters achieves a better fit than that of the model using the optimised values of the top 3 parameters.

Chapter 7

Parameter estimation of tone stack and common-emitter circuits

Parameter estimation here describes the process of retrieving accurate parameter values of physical component models through the identification of a circuit. This objective is motivated by the wide applicability of the results: should the parameters of the constituent components be accurate, these components can be combined in different ways while maintaining an accurate emulation of the circuit's behaviour.

An immediate example of this is the cascading of guitar pedals, where each pedal effectively becomes a subcircuit of the signal chain. Should accurate parameters be available, the changes in input/output impedance when pedals are connected are immediately accommodated for. Recalling the changing of the load of inaccurately estimated parameters in Section 5.3.2, even if the individual pedal's behaviour is reproduced to a high degree of accuracy, unless the parameter values are accurate changing the load of the circuit can introduce significant error.

Two case studies are selected to which parameter estimation is applied: the tone stack from a mid-90s iteration of the Vox AC30 guitar amplifier (the complete schematic of which is shown in Appendix C) , and a generic common-emitter amplifier. The tone stack is assumed to be a linear circuit which alleviates several difficult issues around identifying nonlinear circuits e.g. the higher complexity of nonlinear component models and potential aliasing introduced by distortion. By choosing a linear circuit the objective is to provide a point of comparison for the nonlinear case to qualify the difference in difficulty.

The common-emitter amplifier is again selected – though not the Dallas Rangemaster – to cover the case of a circuit exhibiting nonlinear behaviour that produces pronounced distortion. In Chapter 6, an instance of this circuit was successfully iden-

tified with the results producing a model that captures the input/output behaviour over a range of input signals. The question remains as to whether parameter values can be estimated for the circuit such that the circuit can be modified without introducing error to the corresponding model.

A procedure using elements from Chapter 5 is applied. Results are analysed by directly inspecting the retrieved parameter values, and also by changing the load of the case studies. One notable change is the extension to measurements of a real circuit as opposed to only simulations, which requires considerable additional effort to compensate for non-ideal effects caused by the measurement equipment. This compensation is performed as a part of the initial model design.

7.1 Model design

In contrast to the design of the circuit model used in Chapter 6, the design of the model used for identification here is a heavily involved process. The models must contain suitable components of known value to alleviate possible parameter redundancy as observed in Section 5.1.3.

Potential model failures as observed in the high frequency behaviour of the BC557 Rangemaster (Section 6.3.1) could prevent parameter values from converging to correct values, motivating the inclusion of components that model the non-ideal behaviour in the measurement equipment, such that error between circuit and model can be directly attributed to a failure in the model. Initially, to provide a target for the identification, the parameters are selected and measured.

7.1.1 Parameterisation

As with the calibration of the Rangemaster, the two case studies in this chapter are designed on a breadboard to facilitate direct measurements of component parameters. An LCR meter was purchased to provide accurate measurements of component parameters, the DER DE-5000. By purchasing a new unit, the meter will be within the tolerances specified by the factory calibration, enabling a qualification of the accuracy of the measurements. The selected meter also features a self-calibration feature designed to factor out the capacitance/resistance/inductance of the test probes, further minimising the error introduced to the measurement.

Each linear component was measured immediately prior to the capture of the identification signals, minimising any change in circuit behaviour related to temperature fluctuations that occur over time.

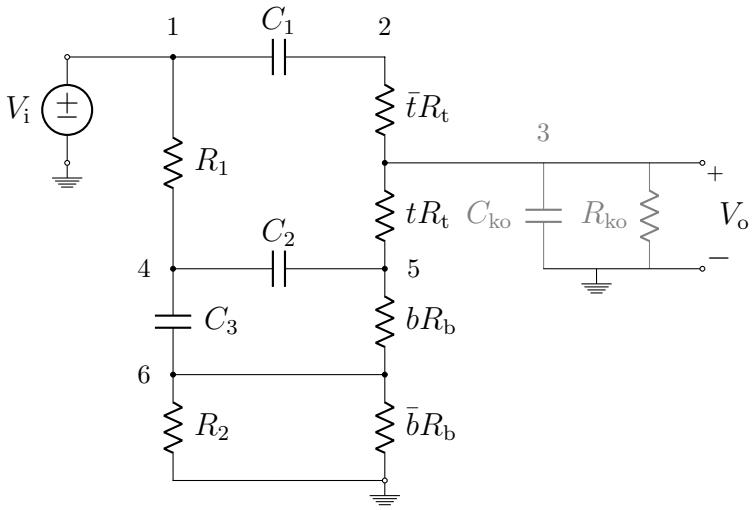


Figure 7.1: Tone stack schematic from the Vox AC30 labelled with nodal indices and known output components C_ko and R_ko.

Tone stack

The schematic of the tone stack from the Vox AC30 amplifier is shown in Figure 7.1. The circuit consists of 4 resistors, 2 of which are potentiometers controlling the treble and bass of the output, and 3 capacitors. Each component is modelled with a single parameter resulting in the vector

$$\boldsymbol{\theta}_\mathrm{ts} = [R_1 \;\; R_\mathrm{t} \;\; R_\mathrm{b} \;\; R_2 \;\; C_1 \;\; C_2 \;\; C_3], \tag{7.1}$$

of $\dim(\boldsymbol{\theta}_\mathrm{ts}) = 7$.

To model the position of the potentiometers two control variables are introduced: t for the treble potentiometer R_t and b for the bass potentiometer R_b. A second variable is defined as $\bar{t} = 1 - t$ to note the remaining resistance on the second resistor used to model the potentiometers. These control variables are not parameters, and are not the target of the identification. Estimating the position of the potentiometer wiper along the track is beyond the scope of this work which is strictly to estimate the component parameters, and therefore the potentiometer is fixed in places that are assumed as known: at the extreme ends of the track. These positions correspond to 0 and 1 for t and b, but inherent resistance in the wire, terminals and wiper of the potentiometer will be non-zero, so these values are approximated for the low end as $t_\mathrm{l} = b_\mathrm{l} = 1 \times 10^{-5}$ and for the high end $t_\mathrm{h} = b_\mathrm{h} = 1 - 1 \times 10^{-5}$.

From the original schematic (shown in Appendix C) the specified parameter values are shown in Table 7.1. Direct measurements of each component are shown with their

Table 7.1: Component parameters of the tone stack, from left to right: component symbol and units, component material, specified parameter value from schematic, directly measured value from LCR meter and tolerance of the measurement.

$\boldsymbol{\theta}_{\mathrm{ts}}$	Unit	Material	Spec.	Direct	Tol.
R_1	kΩ	Metal Film	100	99.35	$\pm 0.52\%$
R_{t}	MΩ	Conductive Plastic	1	1.023	$\pm 1.03\%$
R_{b}	MΩ	Conductive Plastic	1	0.934	$\pm 1.03\%$
R_2	kΩ	Metal Film	10	9.947	$\pm 0.32\%$
C_1	pF	Mica	56	56.84	$\pm 1.28\%$
C_2	nF	Polypropylene	22	21.65	$\pm 0.39\%$
C_3	nF	Polypropylene	22	21.93	$\pm 0.39\%$

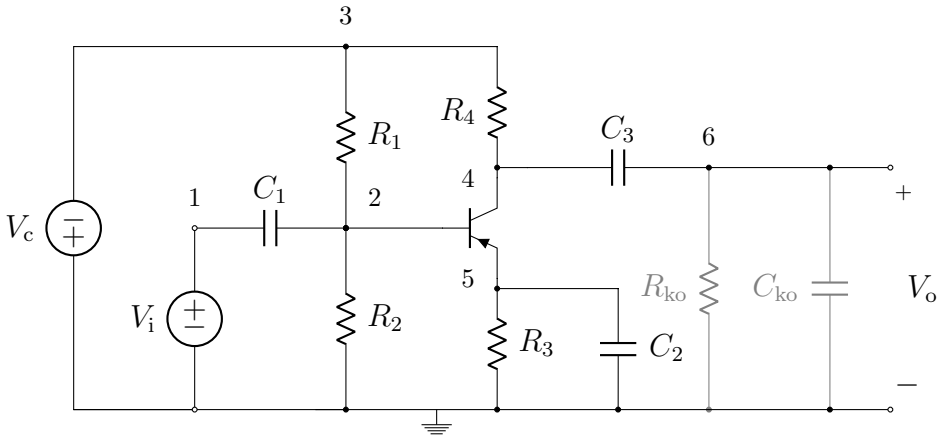


Figure 7.2: Schematic of the common-emitter amplifier labelled with nodal indices and additional known components R_{ko} and C_{ko}.

respective tolerances as derived from the LCR datasheet[1]. Materials are also noted as non-ideal behaviour is typically related to this property. Although the LCR meter does provide the function of measuring secondary parameters (e.g. series resistance with a capacitor), these values have been omitted as the focus of this initial study is to achieve the highest accuracy with only single parameter linear component models.

Common-emitter amplifier

The specific schematic of the common-emitter amplifier identified in this chapter is shown in Figure 7.2, including the specific known components that will be discussed

[1] `https://bholmesqub.github.io/thesis/files/DE-5000.pdf`

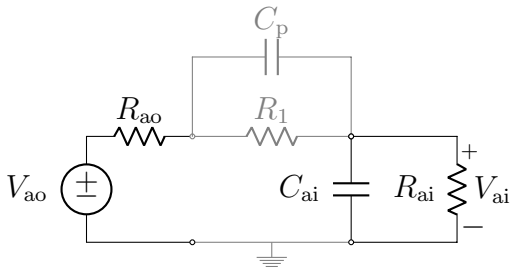


Figure 7.3: An equivalent circuit model of the DAQ used to measure each circuit, relating analogue output V_ao to input V_ai. Grey components mark the series resistance used to estimate the capacitance of the DAQ, and the parasitic capacitance of the breadboard.

in the following sections. The parameter vector $\boldsymbol{\theta}_\mathrm{ce}$ is different from that of the Rangemaster in the previous chapter in that the resistor at the output is assumed to be known, reducing the vector to

$$\boldsymbol{\theta}_\mathrm{ce} = [R_1 \;\; R_2 \;\; R_3 \;\; R_4 \;\; C_1 \;\; C_2 \;\; C_3 \;\; I_\mathrm{s} \;\; N \;\; \beta_\mathrm{f} \;\; \beta_\mathrm{r}] \tag{7.2}$$

which again uses single parameter linear components and the Ebers-Moll BJT model, with $\dim(\boldsymbol{\theta}_\mathrm{ce}) = 11$.

Linear component parameter values were taken from the Rangemaster circuit with one change: C_2 was dropped from $47\,\mu\mathrm{F}$ to $1\,\mu\mathrm{F}$ to enable the selection of a non-polarised capacitor. This decision was made to allow better flexibility for values of V_c, to use an inverted supply voltage may have damaged a polarised capacitor.

Specified and measured values of the linear component parameters are shown in Table 7.2. Again the tolerances of the LCR meter are displayed and secondary parameters are omitted.

The BJT selected for the circuit is the 2N3906 silicon BJT, with the 'specified' parameter values in Table 7.2 taken from LTspice. A silicon BJT was selected for temperature stability, such that if multiple measurements are to be average the self-heating of the BJT will have less of an impact.

7.1.2 Measurement calibration and compensation

The measurement device used in this chapter is the National Instruments USB-6251 DAQ. Specified in the datasheet[2] are 3 parameters for the input and output impedances of the analogue connections: the analogue output has a series resistance $R_\mathrm{ao} = 0.2\,\Omega$, and the analogue input has a parallel impedance of $R_\mathrm{ai} > 10\,\mathrm{G}\Omega$ and $C_\mathrm{ai} = 100\,\mathrm{pF}$.

[2] `https://bholmesqub.github.io/thesis/files/NI_6251.pdf`

Table 7.2: Component parameters of the common-emitter amplifier, from left to right: component symbol and units, component material, specified parameter value from schematic, directly measured value from LCR meter and tolerance of the measurement.

$\boldsymbol{\theta}_{\mathrm{ce}}$	Unit	Material	Spec.	Direct	Tol.
R_1	kΩ	Metal film	470	467.4	$\pm 1.06\%$
R_2	kΩ	Metal film	68	68.16	$\pm 0.52\%$
R_3	kΩ	Carbon film	3.9	3.834	$\pm 0.35\%$
R_4	kΩ	Metal film	10	9.998	$\pm 0.32\%$
C_1	nF	Polypropylene	4.7	4.562	$\pm 0.34\%$
C_2	nF	Ceramic	1000	956.4	$\pm 0.32\%$
C_3	nF	Polypropylene	10	10.202	$\pm 0.32\%$
I_{s}	pA	-	10	-	-
N	-	-	1	-	-
β_{f}	A/A	-	200	-	-
β_{r}	A/A	-	4	-	-

Quickly inspecting the case study circuits it can be seen that the value of R_{ao} will have little effect on the behaviour of both circuits. The minimal impedance to ground of both circuits is at least 4 orders of magnitude higher. The same can be assumed for R_{ai} which is 3 orders of magnitude higher than the largest series impedance. The high series resistance caused by R_{t} in the tone stack may form a low-pass filter with C_{ai}, with a cutoff frequency of approximately $1591\,\mathrm{Hz}$, which would have a notable effect upon the transfer function of the circuit. For this reason, the C_{ai} warrants measurement to compensate for when measuring the case studies.

To measure C_{ai}, a series resistance of $R_1 = 2\,\mathrm{M\Omega}$ was placed on a breadboard between the output and input of the DAQ, as pictured in Figure 7.3. The transfer function was found by driving the circuit with an excitation signal with the parameters $f_{\mathrm{s}} = 1\,\mathrm{MHz}$, $N_{\mathrm{s}} = 1 \times 10^6$, $d_{\mathrm{l}} = 1$ and $d_{\mathrm{h}} = 200 \times 10^3$ which results in a $1\,\mathrm{s}$ signal with frequencies between $1\,\mathrm{Hz} - 200\,\mathrm{kHz}$. The resultant transfer function is displayed in Figure 7.4.

Designing a physical circuit model where the value of R_1 is known and assuming that $R_{\mathrm{ai}} = 10\,\mathrm{G\Omega}$ and $R_{\mathrm{ao}} = 0$ then produces a 1D identification problem, $\theta_{\mathrm{dq}} = C_{\mathrm{ai}}$. During an initial test of the DAQ system the input was connected directly to the output with no load, and a transfer function was taken. The amplitude response of this test

was flat except for uncorrelated noise, but in the high frequency region of the phase response, a systematic curve error was observed. For this reason the data used in the identification was exclusively the amplitude response.

Optimising the model with the Nelder-Mead simplex algorithm as implemented in MATLAB's `fminsearch` [95] with the initial value of $C_{\mathrm{ai}} = 100\,\mathrm{pF}$ produces an optimised value of $C_{\mathrm{ai}} = 56.86\,\mathrm{pF}$. This result is shown in Figure 7.4 by the curve marked RC.

At high frequencies the RC circuit model deviates from the measurement indicating additional effects. One likely component of this is parasitic capacitance of the breadboard, caused by the parallel plates used to connect the components. Adding C_{p} as shown in Figure 7.3 attempts to model this effect, extending the parameters to $\theta_{\mathrm{dq}} = [C_{\mathrm{ai}},\quad C_{\mathrm{p}}]$. Optimising this model using initial values of $C_{\mathrm{ai}} = 56.86\,\mathrm{pF}$ and $C_{\mathrm{p}} = 1\,\mathrm{pF}$ results in $C_{\mathrm{ai}} = 56.25\,\mathrm{pF}$ and $C_{\mathrm{p}} = 680.35\,\mathrm{fF}$.

The results of the RC circuit with the parasitic capacitance is shown again in Figure 7.4. The error in the amplitude response has been reduced over that of the RC without the parasitic capacitance, but additional error has been introduced to the phase response. This indicates that there is behaviour in the circuit that is not being adequately captured.

To determine the exact cause of the behaviour of the circuit would require a significant further time investment in the study and not necessarily provide a useful result. As a first approximation of the measurement equipment C_{ao} will be included in both circuit models, but careful attention must be paid to the high frequency content of the measured signals to ensure that particularly the unidentified behaviour in the phase response does not reduce the accuracy of the identification procedure. The capacitor at the output is present in the schematic of both case studies marked as C_{ko} for the known output capacitance.

7.1.3 Circuit analysis and modelling

Tone stack

The linear tone stack can be modelled using a transfer function as has been previously applied to the RC circuit in Chapter 5. One adaptation is required in that the tone stack features two potentiometers. As mentioned during the discussion of the tone stack's parameterisation, the control variables will each have 2 positions producing in 4 possible combinations of potentiometer positions, resulting in 4 different transfer functions. Each different transfer function is treated as a static snapshot of the behaviour of the

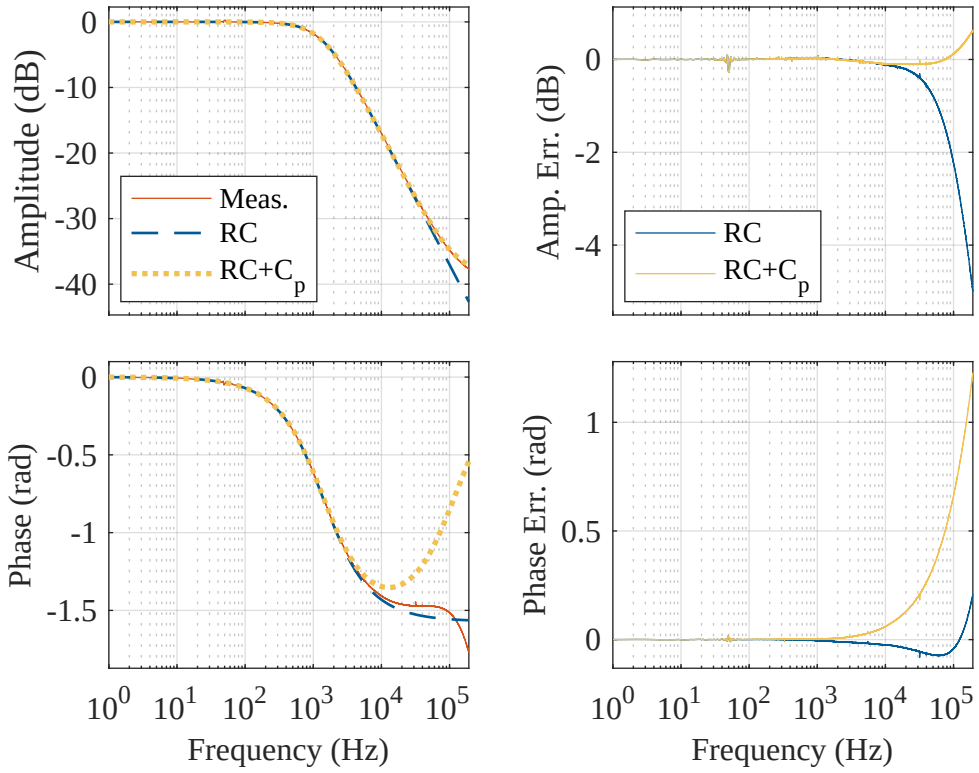


Figure 7.4: Plot of amplitude and phase response of the DAQ with a $2\,\mathrm{M}\Omega$ resistance placed between output and input. Two models have then been fit, and RC circuit and an RC circuit with a parasitic capacitance C_p across the resistor, their error shown in the right plots.

circuit, combined into a single data set as described further in Section 7.2.

A conclusion from Chapter 5 was that by introducing a component of known value into the circuit parameter redundancy could be alleviated. As the capacitance of the DAQ input has been added to the model to compensate for any change in measured behaviour, this known component already exists. However the circuit must still be checked for parameter redundancy.

Visually inspecting the schematic reveals two components of the same impedance in parallel without other connections: R_2 and $\bar{b}R_\mathrm{b}$. Should it be the case that $b = 1$ then parameters R_2 and R_b would not be estimable as they only appear in parallel. Should $b = 0$, R_2, node 6 is shorted to ground and therefore R_2 has no effect over the circuit behaviour making said parameter inestimable. Using multiple values of b will change the value of the combination, ensuring there are at least two different equations for the two parameters. This factor must be remembered during the measurement procedure.

The Laplace form of the transfer function contains 7 terms resultant of many operations upon the constituent parameters. To simplify the parameter redundancy analysis the reduced MNA form can be used. Conductances for the different cases of resistors and potentiometers are defined as $G_\mathrm{R} = 1/R$, $G_\mathrm{Rt} = 1/tR_\mathrm{t}$ and $G_{\mathrm{R}\bar{t}} = 1/\bar{t}R_\mathrm{t}$ (and equivalently for the bass potentiometer). The reduced MNA form is then given by:

$$\begin{bmatrix} G_{\mathrm{R1}} & 0 & 0 & -G_{\mathrm{R1}} & 0 & 0 \\ 0 & G_{\mathrm{R}\bar{t}} & -G_{\mathrm{R}\bar{t}} & 0 & 0 & 0 \\ 0 & -G_{\mathrm{R}\bar{t}} & G_{\mathrm{R}\bar{t}} + G_{\mathrm{Rt}} + G_{\mathrm{Rko}} & 0 & -G_{\mathrm{Rt}} & 0 \\ -G_{\mathrm{R1}} & 0 & 0 & G_{\mathrm{R1}} & 0 & 0 \\ 0 & 0 & -G_{\mathrm{Rt}} & 0 & G_{\mathrm{Rt}} + G_{\mathrm{Rb}} & -G_{\mathrm{Rb}} \\ 0 & 0 & 0 & 0 & -G_{\mathrm{Rb}} & G_{\mathrm{R}\bar{b}} + G_{\mathrm{R2}} + G_{\mathrm{Rb}} \end{bmatrix} \begin{bmatrix} V_\mathrm{i} \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \end{bmatrix} + \begin{bmatrix} C_1 & -C_1 & 0 & 0 & 0 & 0 \\ -C_1 & C_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & C_\mathrm{ko} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_2 + C_3 & -C_2 & -C_3 \\ 0 & 0 & 0 & -C_2 & C_2 & 0 \\ 0 & 0 & 0 & -C_3 & 0 & C_3 \end{bmatrix} \begin{bmatrix} V_\mathrm{i}' \\ V_2' \\ V_3' \\ V_4' \\ V_5' \\ V_6' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{7.3}$$

The exhaustive summary is given by

$$\begin{aligned} \boldsymbol{\kappa}(\boldsymbol{\theta}_\mathrm{ts}) = [&G_{\mathrm{R1}},\ \ G_{\mathrm{R}\bar{t}},\ \ G_{\mathrm{R}\bar{t}} + G_{\mathrm{Rt}},\ \ G_{\mathrm{Rt}},\ \ G_{\mathrm{Rt}} + G_{\mathrm{Rb}}\ \ G_{\mathrm{Rb}}, \\ &G_{\mathrm{R}\bar{b}} + G_{\mathrm{R2}} + G_{\mathrm{Rb}},\ \ C_1,\ \ C_2,\ \ C_2 + C_2,\ \ C_3], \end{aligned} \tag{7.4}$$

which upon testing yields the $\mathrm{rank}(\mathbf{J}(\boldsymbol{\kappa})) = 7$ and therefore each parameter is estimable. Parameter redundancy was revealed in the previous chapter by dividing the

whole model by one parameter to produce terms without any parameters. In this case as a C_{ko} and R_{ko} are already included, any operation to remove a parameter would combine with these known values and as such no parameter redundancy is present in the model.

The redundancy analysis is then repeated using upon the form of the model that will be used in the simulation of the circuit, in this case, a transfer function. The transfer function of the tone stack is of the form

$$H(s) = \frac{b_3 s^3 + b_2 s^2 + b_1 s}{a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + 1}, \tag{7.5}$$

where coefficients have been normalised by a_0 to remove one coefficient from the vector thus maximally simplifying the transfer function. This has a corresponding exhaustive summary of

$$\hat{\boldsymbol{\kappa}}(\boldsymbol{\theta}_{\mathrm{ts}}) = [b_3,\ \ b_2,\ \ b_1,\ \ a_4,\ \ a_3,\ \ a_2,\ \ a_1]. \tag{7.6}$$

Immediately it is clear that it is possible for the model to have estimable parameters as the length of the exhaustive summary is the same as that of the parameter vector. The full coefficients are omitted due to their complexity. Inspecting the determinant of the Jacobian results in $\mathrm{rank}(\mathbf{J}(\hat{\boldsymbol{\kappa}})) = 7$ confirming that each parameter is estimable using the transfer function model.

Having compensated for the measurement equipment by setting $C_{\mathrm{ko}} = C_{\mathrm{ai}}$ and verifying that each parameter is theoretically estimable, the model is now suitable to test with simulated measurements to determine whether the parameter values can be retrieved from an ideal identification scenario.

Common-emitter amplifier

A state-space model is selected to model the common-emitter amplifier. The volume potentiometer in the previously modelled common-emitter amplifiers is treated as fixed at maximum volume and therefore needs no further consideration as required for the tone stack. To ensure the model is suitable for identification all that remains is the verification that each parameter is estimable.

No obvious redundancy is clear from visually inspecting the schematic, so the

model can be checked symbolically. The reduced MNA model is given by

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & G_{\mathrm{R1}}+G_{\mathrm{R2}} & -G_{\mathrm{R1}} & 0 & 0 & 0 \\ 0 & -G_{\mathrm{R1}} & G_{\mathrm{R1}}+G_{\mathrm{R4}} & -G_{\mathrm{R4}} & 0 & 0 \\ 0 & 0 & -G_{\mathrm{R4}} & G_{\mathrm{R4}} & 0 & 0 \\ 0 & 0 & 0 & 0 & G_{\mathrm{R3}} & 0 \\ 0 & 0 & 0 & 0 & 0 & G_{\mathrm{Rko}} \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}} \\ V_2 \\ V_{\mathrm{c}} \\ V_4 \\ V_5 \\ V_6 \end{bmatrix} + \begin{bmatrix} C_1 & -C_1 & 0 & 0 & 0 & 0 \\ -C_1 & C_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & C_3 & 0 & -C_3 \\ 0 & 0 & 0 & 0 & C_2 & 0 \\ 0 & 0 & 0 & -C_3 & 0 & C_3+C_{\mathrm{ko}} \end{bmatrix} \begin{bmatrix} V_{\mathrm{i}}' \\ V_2' \\ V_{\mathrm{c}}' \\ V_4' \\ V_5' \\ V_6' \end{bmatrix} = \begin{bmatrix} 0 \\ -I_{\mathrm{b}} \\ 0 \\ -I_{\mathrm{c}} \\ I_{\mathrm{b}}+I_{\mathrm{c}} \\ 0 \end{bmatrix}, \tag{7.7}$$

with a corresponding exhaustive summary of

$$\begin{aligned} \boldsymbol{\kappa}(\boldsymbol{\theta}_{\mathrm{ce}}) = [&G_{\mathrm{R1}}+G_{\mathrm{R2}},\ G_{\mathrm{R1}},\ G_{\mathrm{R1}}+G_{\mathrm{R4}},\ G_{\mathrm{R4}},\ G_{\mathrm{R3}},\ C_1,\ C_3,\ C_2, \\ &I_{\mathrm{s}}/\beta_{\mathrm{f}},\ I_{\mathrm{s}}/\beta_{\mathrm{r}},\ I_{\mathrm{s}},\ I_{\mathrm{s}}(\beta_{\mathrm{r}}+1)/\beta_{\mathrm{r}},\ N] \end{aligned}. \tag{7.8}$$

Testing the rank of the Jacobian reveals $\mathrm{rank}(\mathbf{J}(\boldsymbol{\kappa})) = 12$, indicating that each parameter is estimable. Again, with the inclusion of a known component if the model is scaled by a parameter, that parameter combines with the known component and as such there is no parameter redundancy to be revealed. This can be further investigated on the state-space model. To simplify the analysis let the coefficients of the Ebers-Moll exponentials be denoted in matrix form as

$$\mathbf{L} = \begin{bmatrix} I_{\mathrm{s}}/\beta_{\mathrm{f}} & I_{\mathrm{s}}/\beta_{\mathrm{r}} \\ I_{\mathrm{s}} & I_{\mathrm{s}}(\beta_{\mathrm{r}}+1)/\beta_{\mathrm{r}} \end{bmatrix}, \tag{7.9}$$

such that

$$\hat{\boldsymbol{\kappa}}(\boldsymbol{\theta}_{\mathrm{ce}}) = [\mathbf{A},\ \mathbf{B},\ \mathbf{CL},\ \mathbf{D},\ \mathbf{E},\ \mathbf{FL},\ \mathbf{G},\ \mathbf{H},\ \mathbf{KL}, N]. \tag{7.10}$$

All zero value entries can be removed as with duplicate terms. Testing the rank of Jacobian reveals $\mathrm{rank}(\mathbf{J}(\hat{\boldsymbol{\kappa}})) = 12$, confirming that each parameter is estimable for this state-space model of the common-emitter amplifier.

Having determined that the common-emitter amplifier has theoretically estimable parameters and features compensation of the measurement equipment, the model is now suitable for testing with simulated measurements.

7.2 Optimisation formulation

For both case studies the Nelder-Mead simplex method is used to optimise the model with the objective function as defined in (5.37). As the data sets are made up of multiple signals/transfer functions, the definition is extended such that

$$\xi(\hat{\boldsymbol{\theta}}) = \sum_{m=1}^{M} \left(\frac{1}{\eta_m} \sum_{n=0}^{N_\mathrm{s}-1} \epsilon_m(n, \hat{\boldsymbol{\theta}})^2 \right), \tag{7.11}$$

where m is the index of the signal/transfer function within a set of M signals/transfer functions.

The central challenge in designing a successful identification procedure lies in the capture of sufficient behaviour in the excitation signal such that the final estimated values of each parameter are accurate with respect to the directly measured values. With the case studies in Chapter 5, the number of parameters was low enough that only a coarse design was required to produce excellent results, but as the case studies in this chapter have more parameters the challenge is increased.

7.2.1 Excitation signal

Tone stack

The excitation signal for the tone stack was designed to provide a set of transfer functions that could be directly compared with the model as opposed to time domain signals. To maximise resolution a high sample rate was chosen, $f_\mathrm{s} = 1\,\mathrm{MHz}$, and a $1\,\mathrm{s}$ signal was generated with $N_\mathrm{s} = 1 \times 10^6$ so that the minimum captured frequency was $1\,\mathrm{Hz}$. Empirically, it was found that including frequencies close to DC improved the convergence of parameter values, motivating the choice of capturing the behaviour over 4 octaves beneath the commonly stated limit of human hearing. Values of $d_\mathrm{l} = 1$ and $d_\mathrm{h} = 20 \times 10^3$ defined a set of frequencies from $1\,\mathrm{Hz} - 20\,\mathrm{kHz}$. A peak voltage was chosen to correspond with the amplitude used by the LCR meter, $V_\mathrm{p} = 0.5\,\mathrm{V}$. The inverse of the simulated transfer functions was applied as a frequency amplitude envelope to maximise SNR across the measured range.

At the time of the research, the values of A_d used in the multi-sine excitation signal were not modified to ensure that their sum is equal to 1. Instead they were normalised using $N_\mathrm{s}/\sqrt{d_\mathrm{h} - d_\mathrm{l}}$, which provided a signal with a standard deviation that is constant and independent of the signal parameters. If $\sum_{d=d_\mathrm{l}}^{d_\mathrm{h}} A_d \neq 1$, then there are scenarios in which a high crest-factor can arise, but for the excitation signal used in this work

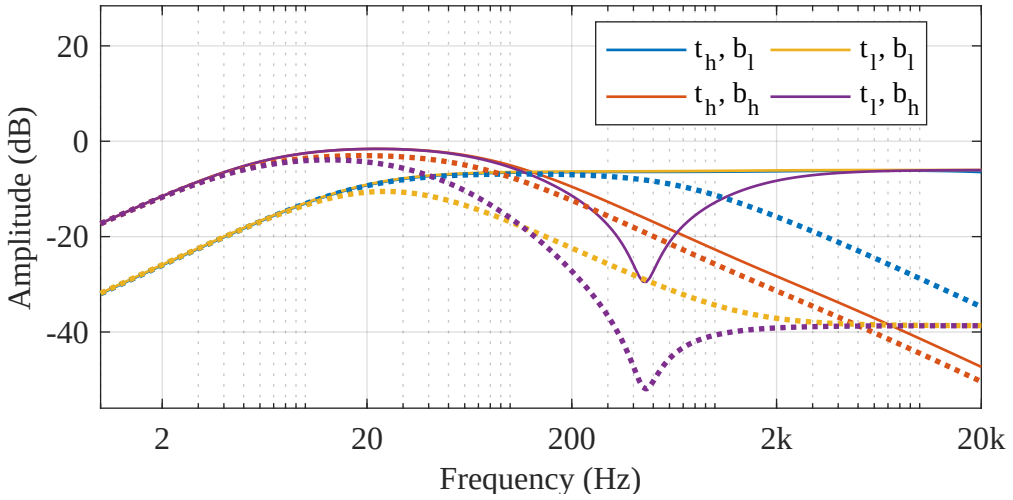


Figure 7.5: Amplitude responses of the different potentiometer positions, subscript h marking a position close to 1 and l close to 0. The solid line responses use $C_{\mathrm{ko}} = 56\,\mathrm{pF}$ and dashed lines use $C_{\mathrm{ko}} = 56\,\mathrm{pF} + 4.7\,\mathrm{nF}$.

– prior to the application of the Hann window – the crest factor was 3.36 indicating a low peak-to-peak voltage.

As with the excitation signal in Chapter 6, amplitudes were not chosen to satisfy $\sum_{d=d_\mathrm{l}}^{d_\mathrm{h}} A_d = 1$, but instead scaled using $N_\mathrm{s}/\sqrt{d_\mathrm{l} - d_\mathrm{h}}$. The excitation signal for the tone stack has a maximum crest factor of 4.7592 and 2.7843, with and without the inverse transfer function weighting, respectively.

To mitigate potential high-frequency error observed in the DAQ, two adaptations were made. The first is that the objective function uses only the magnitude of the transfer function, such that in (5.37) and (5.38) $y = |H(j\omega_n)|$. By only utilising the amplitude response, the disparity between high frequency amplitude and phase responses observed in the measurement equipment is circumvented. In addition to this, a capacitor of value $4.7\,\mathrm{nF}$ was placed across the output terminal, resulting in $C_{\mathrm{ko}} = C_{\mathrm{ao}} + 4.7\,\mathrm{nF}$. The increase in capacitance reduces high frequency amplitude in the transfer function when $t = t_\mathrm{l}$, which will reduce the amount of potential error in the objective function. Figure 7.5 shows the difference in simulated amplitude response between using just the capacitance of the DAQ and using the additional output capacitor, reducing high frequency amplitude for each transfer function.

Additional samples were removed from the tone stack transfer functions, with samples above $20\,\mathrm{kHz}$ removed and the resulting transfer function then downsampled such that only 256 points remained, spaced logarithmically over the defined frequency range. For the transfer function of the measured tone stack circuits a further 6 samples

were removed, 4 from around $50\,\mathrm{Hz}$ and 2 more from $100\,\mathrm{Hz}$ due to the high amount of noise likely caused by proximity to mains power circuits.

Common-emitter amplifier

In a similar manner to the 4 combinations of potentiometer positions for the tone stack, 4 different excitation signals were designed for the common-emitter amplifier. Each signal was designed to expose different behaviour of the amplifier to provide options when identifying the model using simulated measurements: should one signal not be sufficient a combination can be used.

Table 7.3 show the signal parameters for the 4 different excitation signals. The simulated output measurements of the common-emitter amplifier are displayed in Figure 7.6.

Two major differences exist between the set of signals: two different supply voltages are used, and for each different supply voltage one signal has a low amplitude and wide bandwidth, and the other a high amplitude and narrow bandwidth. Two different supply voltages were used to change the bias of the BJT, trying to better expose it's behaviour over a range of operating conditions. The other differences were chosen to capture different aspects of the circuits behaviour.

Selecting a low amplitude and wide bandwidth signal provides a result akin to a transfer function. An attempt was made at measuring the transfer function of the circuit, but due to it's distortion and high noise floor an excitation could not be designed that captured a transfer function, in the same way as the tone stack was measured. The high amplitude signal was chosen to ensure the distortion of the circuit was captured as it is a critical sonic characteristic, as well as providing significant information about the behaviour of the BJT. For this signal, the bandwidth was reduced to prevent aliasing in the measurement or model (the selected DAQ features no anti-aliasing filter). Signal length was reduced to compensate for the high sample frequency required to reduce aliasing.

Optimal performance was found from the DAQ at $f_\mathrm{s} = 1\,\mathrm{MHz}$ but this sample rate is not necessarily ideal for the identification, so downsampling was performed on the low amplitude signals which had lower bandwidth due to less high frequency distortion resulting in $\hat{f}_\mathrm{s}$. To compensate for high frequency discrepancies between circuit and model, and to a lesser extent low frequency noise, output signals were filtered with a bandpass filter consisting of both a 2nd order Butterworth low pass and high pass filter, the high pass cutoff frequency noted by $\omega_{\mathrm{c}1}$ and the low pass cutoff noted by $\omega_{\mathrm{c}2}$. Both filters were applied in the objective function such that $y = l(V_\mathrm{o})$ where l is the

Table 7.3: Specification of the signals used for the Common-Emitter amplifier.

#	$\hat{f}_\mathrm{s}$ (MHz)	N_s	Dur. (s)	V_c (V)	f_l (Hz)	f_u (kHz)	d_l	d_u	V_p (V)	$\omega_{\mathrm{c}1}$ (Hz)	$\omega_{\mathrm{c}2}$ (kHz)
1	1	20×10^3	0.02	-9	50	2	10	400	3	50	4
2	0.1	100×10^3	1	-9	1	20	1	20×10^3	0.6	1	30
3	1	20×10^3	0.02	-5	50	2	10	400	3	50	4
4	0.1	100×10^3	1	-5	1	20	1	20×10^3	0.6	1	30

bandpass filter.

In the calibration of the Rangemaster, a windowed signal was used to ensure different signal levels were represented. As a combination of signals with different amplitudes is possible for this identification task, a window is not necessary and therefore not applied.

Weighting was applied to the sine components to produce a flat output amplitude response. For the low amplitude signals, this was simply the inverse transfer function found from the model using the measured linear component values and specified BJT values. For the high amplitude signals, the inverse transfer function was weighted and scaled such that

$$A_\mathrm{d} = 1 + \left(1 - 10^{-4}\right)\left(\left|\hat{H}(j\omega_\mathrm{d})^{-1}\right| - 1\right), \tag{7.12}$$

where $\hat{H}(j\omega)^{-1}$ is the inverse transfer function normalised such that its largest magnitude is 1. Scaling the inverse transfer function prevents low frequencies from dominating the signal which as the signal is a higher amplitude, could cause distortion products to overwhelm other behaviour of the circuit at higher frequencies. The amplitudes were then scaled by $N_\mathrm{s}/\sqrt{d_\mathrm{l} - d_\mathrm{h}}$ resulting in a maximum crest factor of 2.4381 and 2.7210 with and without the application of the inverse transfer function respectively.

7.2.2 Data set selection

Having designed a dataset consisting of candidate excitation signals and their corresponding output/transfer function measurements, it must then be determined what combination of data provides the optimal identification results. The criteria for a successful identification are not only retrieving parameter values, but being able to repeatedly retrieve the parameter values for different initial parameter value sets.

Simulated measurements of the different combinations are used to ascertain which combination will provide the lowest values of mean parameter error $\langle \epsilon_\theta \rangle$ for both models. By discovering the optimal combination prior to measurements, it is possible that

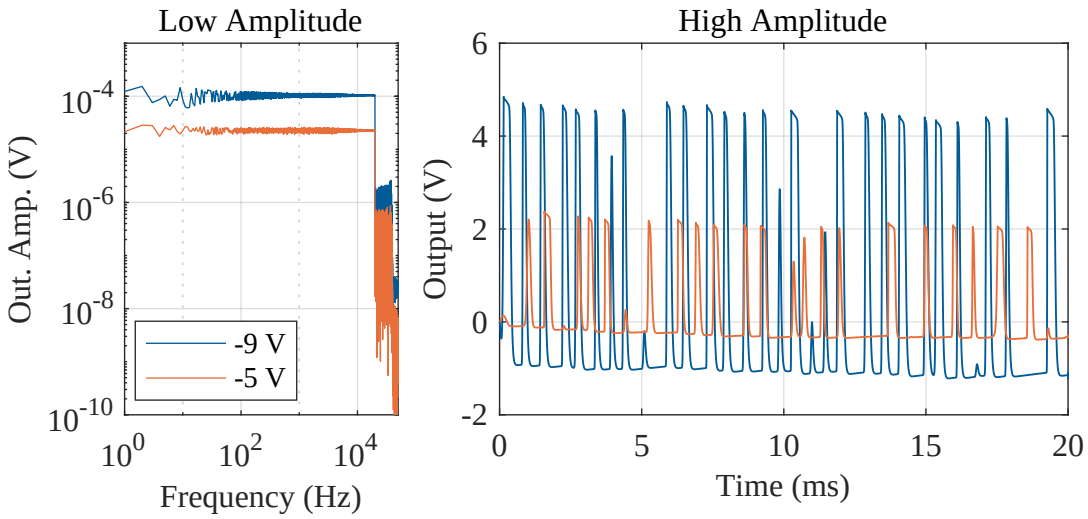


Figure 7.6: Simulated measurements of the common-emitter amplifier. The left plot shows the amplitude spectrum of the low amplitude signals 2 & 4, the right plot the high amplitude signals 1 & 3 in the time domain.

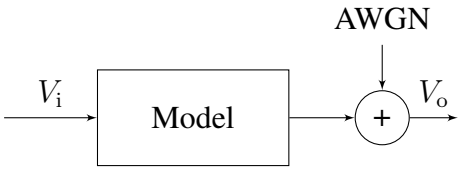


Figure 7.7: Signal diagram depicting how AWGN is applied to simulated measurements.

some signals will not require measuring, reducing the work required to identify each circuit. Simulated measurements have been shown in Chapter 5 to achieve a high degree of accuracy when estimating parameter values, likely more than is achievable when working with measured signals. To better predict the parameter accuracy when using real measurements, Additive White Gaussian Noise (AWGN) is applied to all simulated measurements as depicted in Figure 7.7. By including a non-ideal feature of the measurement, the objective is to better approximate a real measurement which will thus yield a better prediction as to the resultant values of $\langle \epsilon_\theta \rangle$. The prediction of parameter accuracy is validated through comparison to actual measurements in the following section.

Tone stack

Four transfer functions exist for the tone stack as defined by the possible combinations of potentiometer positions. Noted during the model design of the tone stack, should

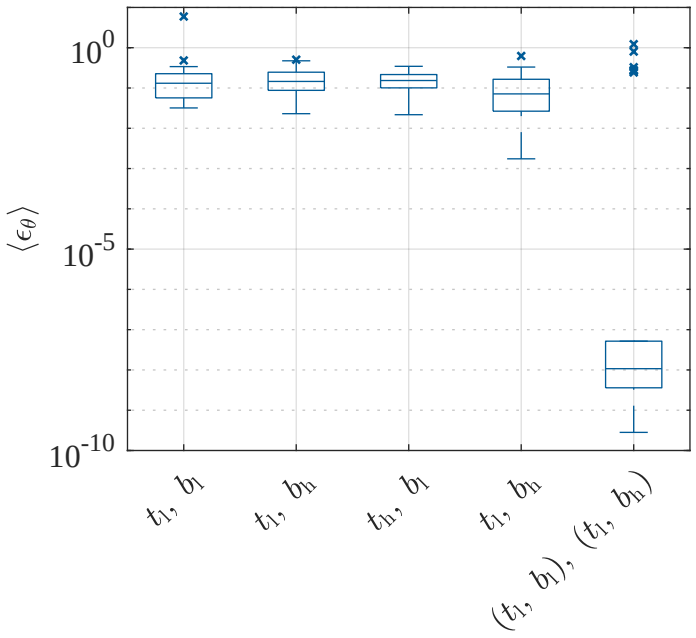


Figure 7.8: Box plot demonstrating the inability of to retrieve tone stack parameters when only using one transfer function, comparing each of the four possible single transfer functions with one case using two transfer functions. 30 optimisations are performed on each combination of transfer functions.

the potentiometer control b only be in one position, possible potentiometer redundancy is introduced. Figure 7.8 demonstrates this redundancy through optimising the model using each single transfer function, and one combination using two transfer functions. No AWGN was used to enable minimal values of $\langle \epsilon_\theta \rangle$. Exit conditions were specified as a change in ξ or θ_{ts} of less than 1×10^{-8}, and for each combination the optimisation was repeated 30 times. Median values of each of the single combinations are around 10%, whereas when two different sets of potentiometer positions are used the median drops to $1 \times 10^{-6}\%$, much higher accuracy than achievable with a measurement from the LCR meter. Evidently parameter values can be retrieved in an ideal situation, but require more than one combination of potentiometer positions for this tone stack.

The remaining task is to predict the optimal set of transfer functions to find minimal values of $\langle \epsilon_\theta \rangle$ that can be achieved repeatedly. The AWGN applied is characterised with a zero-value mean and a variance defined using the 'Random Noise' standard deviation specification of the analogue input of the DAQ. To reduce the effect of the AWGN averaging is employed: the variance of the AWGN is inversely proportional to the number of averages. The chosen analogue input range of the DAQ is $\pm 0.5\,\text{V}$ which is specified to have noise with a standard deviation of $\sigma_{\text{dq}} = 21\,\mu\text{V}$, and 60 averages were used resulting in $\sigma^2 = \frac{21^2}{60}\,\mu\text{V} = 7.35\,\mu\text{V}$.

The results of 30 optimisations performed on the simulated measurements for the combinations of multiple transfer functions of the tone stack are shown in Figure 7.9. The combination with the lowest mean is that using all of the transfer functions as might be expected. Values of $\langle \epsilon_\theta \rangle$ appear to be limited by the amount of noise in the system, with the trend appearing to converge upon a lower limit. From this result it is clear that using all of the transfer function is likely to provide the best results when optimising on measured data. The accuracy of this result and the other data trends will be investigated by comparing the results with the equivalent results of the measured data in the following section.

Common-emitter amplifier

To determine which signals should be used to properly expose the behaviour of the common-emitter amplifier a similar analysis is performed as with the tone stack. An additional factor must be addressed before performing the combination analysis: aliasing. As the common-emitter amplifier model is nonlinear and exhibits significant distortion, aliasing is bound to be present, even if reduced using anti-aliasing techniques. When fitting the model to the circuit measurements, the aliasing behaviour is unlikely to match exactly due to the processing present in the analogue input of the DAQ. If

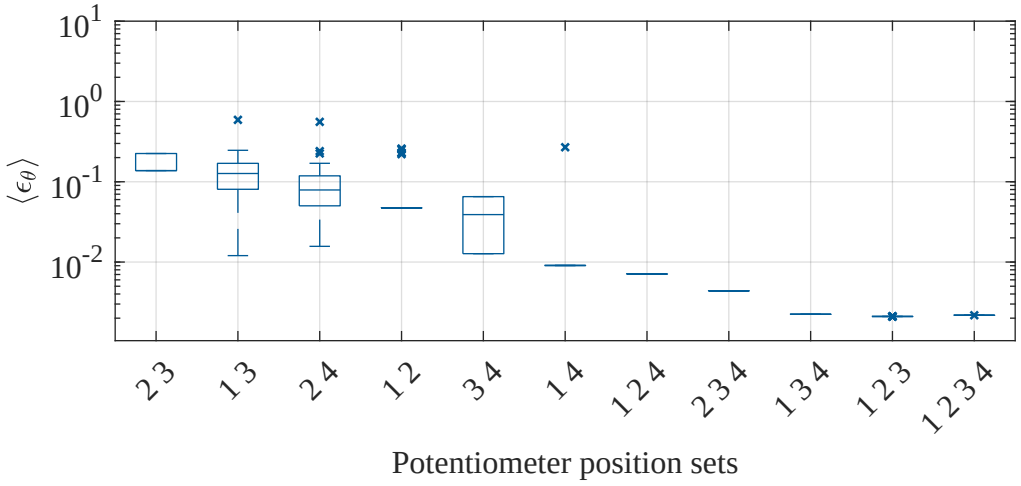


Figure 7.9: Box plot of optimised parameter error for each theoretically estimable combination of potentiometer positions for the tone stack. Combinations are noted by 1: $t_\mathrm{l},\ b_\mathrm{l}$, 2: $t_\mathrm{l},\ b_\mathrm{h}$, 3: $t_\mathrm{h},\ b_\mathrm{l}$, 4: $t_\mathrm{h},\ b_\mathrm{h}$. 30 optimisations were performed for each combination, the resultant values were averaged across parameters and repeats. Results are sorted by number of combinations then by median value of $\langle \epsilon_\theta \rangle$.

fitting the model to simulated measurements and the signals are sampled at exactly the same rate and processed identically, it is possible for the aliasing artefacts to perfectly match, achieving a level of accuracy not possible with measurements.

To overcome this problem, the simulated measurements can be generated at an oversampled sample rate and then downsampled, preventing a perfect fit of aliasing behaviour. Figure 7.10 shows the results of 10 optimisations of the common-emitter amplifier using the high amplitude input/output signals (those with the highest distortion), oversampled at five different levels between $1\times$ and $16\times$ before being downsampled to the original specified sample rate for optimisation. The AWGN is omitted from the outputs to ensure that the accuracy of $\langle \epsilon_\theta \rangle$ is only relative to the oversampling. Without oversampling, the parameters can be retrieved to a high degree of accuracy. Values of both $\langle \epsilon_\theta \rangle$ and ξ quickly increase to a plateau when oversampling is applied, indicating that matching aliasing behaviour is required to best retrieve parameter values.

In the prediction of $\langle \epsilon_\theta \rangle$ for measurements, to ensure a minimal fit from aliasing artefacts, $16\times$ oversampling was selected for the generation of each signal in the simulated measurements. AWGN is applied, 30 averages are used and the input ranges used are different, for the low amplitude measurement $\pm 1\,\mathrm{V}$ is used with a standard deviation of $\sigma_\mathrm{d} = 32\,\mu\mathrm{V}$, resulting in $\sigma^2 = \frac{32^2}{30}\,\mu\mathrm{V} = 34.1\,\mu\mathrm{V}$. For the high amplitude

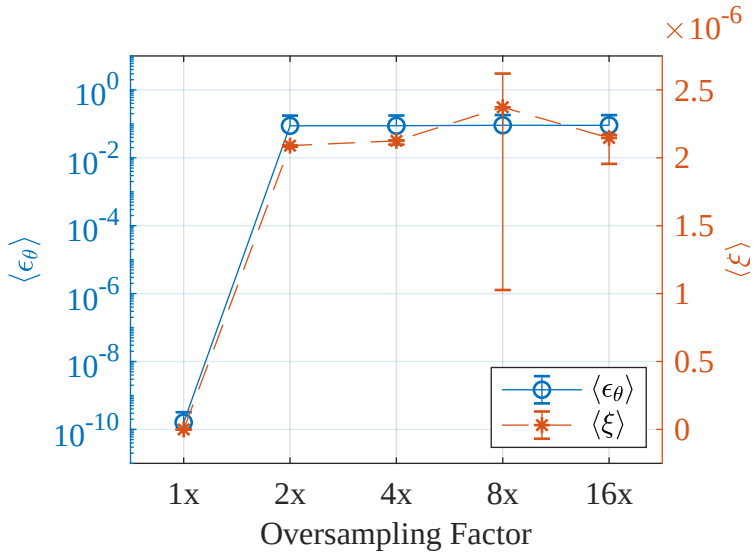


Figure 7.10: Oversampling vs optimised parameter error $\langle \epsilon_\theta \rangle$ and objective function value ξ for the common-emitter amplifier. High amplitude signals 1 & 3 were used in the optimisation, of which there were 10 repeats for each level of oversampling.

the range was set at $\pm 5\,\mathrm{V}$ which has noise with a standard deviation of $\sigma_\mathrm{d} = 140\,\mu\mathrm{V}$, resulting in $\sigma^2 = \frac{140^2}{30}\ \mu\mathrm{V} = 653.3\ \mu\mathrm{V}$. The parameters used to find $\langle \epsilon_\theta \rangle$ are only from the linear components: no accurate parameters are available for the BJT so to predict values of $\langle \epsilon_\theta \rangle$ for measurements, they must also be excluded for the simulated measurement comparison.

Figure 7.11 shows the results of different combinations of the simulated measurements. A similar trend as to the tone stack is observed here with a smaller variance observed as more signals are included, and generally a lower median value. Two results that are anomalous to this trend are found in the sets consisting of only 1, and 1 & 3 signals, these signals are the high amplitude signals, for 1, $V_\mathrm{c} = -9\,\mathrm{V}$ and for 3, $V_\mathrm{c} = -5\,\mathrm{V}$. Results from optimising only using signal 1 results in a median similar to that when using all signals, but with a larger variance including results with lower $\langle \epsilon_\theta \rangle$. Using 1 & 3 results in a low variance and median value lower than any other result. To investigate the difference between optimising using all signals and only using 1 & 3, both will be compared when optimising on measurements from the real circuit.

7.3 Results and validation

Three areas of results are presented to qualify the success of the parameter estimation of both case studies. The first area is in how well the simulated measurements predicted

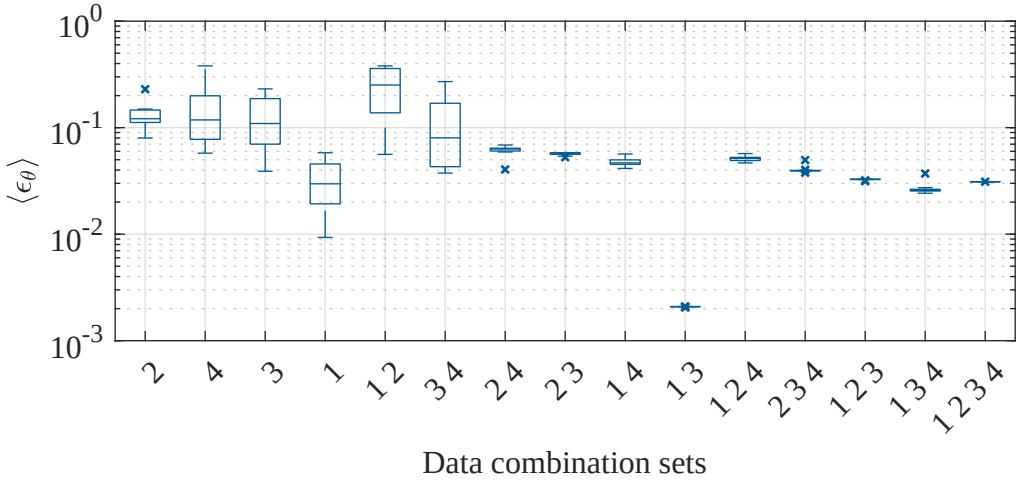


Figure 7.11: Box plot of optimised parameter error of 15 repeats for each possible combination of excitation signals for the common-emitter amplifier. Indices refer to signal number in Table 7.3: 1 & 2 use $V_c = -9\,\mathrm{V}$, 3 & 4 use $V_c = -5\,\mathrm{V}$, odd numbers are high amplitude, even numbers are low amplitude. Box plots are sorted by median value.

the behaviour of the circuit measurements with regards to the optimised parameter error for each combination of signals. Following this an analysis of the optimised parameter values found using the chosen combination(s) is performed, determining the accuracy and reliability of the parameter estimation process. Finally, the models using optimised parameter values are validated by changing the load of the circuit and model, and checking to see whether additional error is introduced.

Though not discussed within this section, sound examples are provided for readers to assess the subjective quality of the optimised models, available online[3].

7.3.1 Validation of combination selection

Tone stack

The exact same procedure as was used to generate the results in Figure 7.9 was repeated, this time using the measured data from the circuit. Figure 7.12 shows the results for the measured data, sorted using the order from Figure 7.9. Comparing the results, the same trend can be observed that using more combinations generally produces lower values of $\langle \epsilon_\theta \rangle$ with less variance.

[3]https://bholmesqub.github.io/thesis/chapters/parameter-estimation/sound-examples/

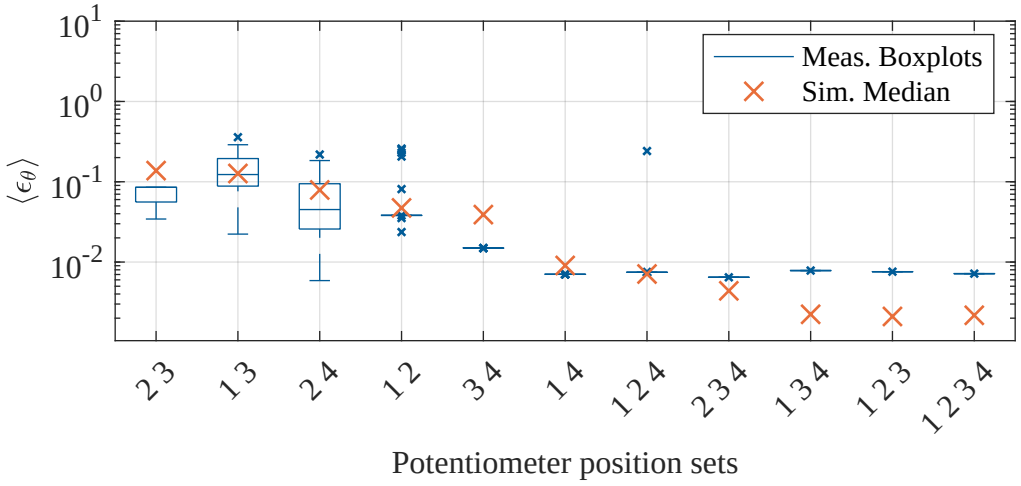


Figure 7.12: Box plot of optimised parameter error for each possible combination of potentiometer positions for the tone stack. Combinations are noted by 1: t_l, b_l, 2: t_l, b_h, 3: t_h, b_l, 4: t_h, b_h. 30 optimisations were performed for each combination, the resultant values were averaged across parameters and repeats. Results are sorted to match the results of Figure 7.9 to enable easy comparison.

The most significant disparity is that for the combinations with the lowest values of $\langle \epsilon_\theta \rangle$, the median values for simulated measurements are approximately five times lower than those of the real measurements. This error indicates that the noise modelling introduced does not sufficiently capture the various sources of error in the circuit to predict the results of the parameter estimation to a high degree of accuracy. This may be due to error such as the systematic high-frequency error observed in Section 7.1.2, or other sources such as $50\,\mathrm{Hz}$ noise transmitted from the mains power sources.

Common-emitter amplifier

The measured counterpart of the signal combination comparison is displayed in Figure 7.13 with the median of the simulated comparison overlayed. The anticipated trend of reducing median value and variance as the number of signals used in the optimisation is observed. The anomalous result of using only signal number 1 yielding lower parameter error than all signals is not repeated in the measured data, though the combination of 1 & 3 does yield a similar median parameter error as when using all signals.

The two anamolous results in the simulated measurements are evidence of a significant difference between model and measurement. From the investigation to the DAQ behaviour, noteable high frequency phase error was noted, but not compensated for

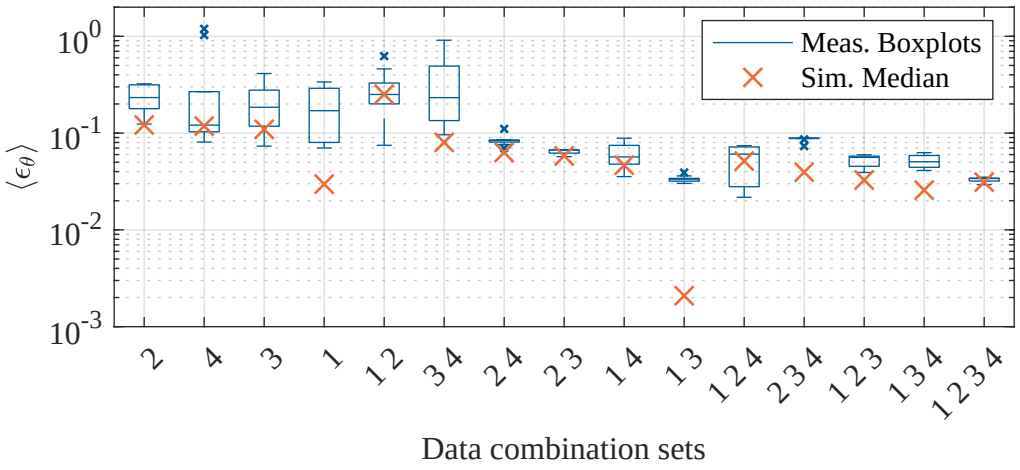


Figure 7.13: Box plot of optimised parameter error of 10 repeats for each possible combination of excitation signals for the common-emitter amplifier. Indices refer to signal number in Table 7.3: 1 & 2 use $V_c = -9\,\mathrm{V}$, 3 & 4 use $V_c = -5\,\mathrm{V}$, odd numbers are high amplitude, even numbers are low amplitude. Box plots are sorted to follow the same order as Figure 7.11.

as the source was unclear. In the transfer function of the tone stack, this was avoided through ignoring the phase response of the circuit. For the common-emitter amplifier, it is possible that the high frequency transitions in the signal are misaligned due to this phase error, preventing a high level of fit. Other causes may be differences in distorting behaviour of the BJT between model and device, as found in Chapter 4 albeit for germanium BJTs, or variation in aliasing between model and measurements.

These two anomalous results present more of an issue than the general mismatch found when comparing simulated and real measurement combinations of the tone stack. This inaccuracy could be attributed to one of several issues, potentially a difference in noise characteristics between simulation and measurement that the low amplitude signals may be more sensitive to, or simply that the model can fit better to the simulated measurements than the behaviour of the real circuit.

7.3.2 Identified parameters

Tone stack

Using all of the possible transfer functions of the tone stack, the parameters of the circuit were estimated. To provide some understanding as to the repeatability of the

Table 7.4: Parameters of the tone-stack, from left to right: component symbol and units, the value specified by the schematic, the directly measured value using an LCR meter and its tolerance, the resulting θ_{ts} with the minimum ξ from the 100 sets, its error relative to the directly measured values, and the standard deviation of the full set.

$\boldsymbol{\theta}_{\mathrm{ts}}$	Spec.	Direct	Tol.	Estim.	Error	RSD
R_1 (kΩ)	100	99.35	$\pm 0.52\%$	100.21	0.867%	$1.22 \times 10^{-4}\%$
R_{t} (MΩ)	1	1.023	$\pm 1.03\%$	1.0213	0.170%	$6.28 \times 10^{-5}\%$
R_{b} (MΩ)	1	0.934	$\pm 1.03\%$	0.93910	0.535%	$1.88 \times 10^{-4}\%$
R_2 (kΩ)	10	9.947	$\pm 0.32\%$	10.032	0.856%	$1.29 \times 10^{-4}\%$
C_1 (pF)	56	56.84	$\pm 1.28\%$	58.161	2.32%	$7.38 \times 10^{-5}\%$
C_2 (nF)	22	21.65	$\pm 0.39\%$	21.670	0.0923%	$1.54 \times 10^{-4}\%$
C_3 (nF)	22	21.93	$\pm 0.39\%$	21.969	0.178%	$1.89 \times 10^{-4}\%$

parameter estimation, 100 repeats were performed using starting parameter values selected from $\pm 40\%$ of the accurate parameter values. The results of the optimisations are displayed in Table 7.4.

From the 100 final sets of $\boldsymbol{\theta}_{\mathrm{ts}}$, one was selected which had the minimal corresponding value of ξ, noted by the column header 'Estim.' which is used as a benchmark to compare to the directly measured parameter values. Largely the parameters were within 1% of the directly measured value, the exception being C_1. Further, the observed error is less than the tolerance of the LCR meter used to directly measure each parameter value for 4 out of 7 parameters (R_{t}, R_{b}, C_2, and C_3).

The Relative Standard Deviation (RSD) of the full set of optimised $\boldsymbol{\theta}_{\mathrm{ts}}$ is shown to indicate the repeatability of the results. The maximum RSD value found is 1.89×10^{-4}, revealing that from a wide variety of initial parameter values a highly accurate result can be achieved.

Measurements of the tone stack transfer function and a model using the selected $\boldsymbol{\theta}_{\mathrm{ts}}$ are illustrated in Figure 7.14. The trends of each transfer function have been matched for both amplitude and phase responses, with the exception of the high frequency phase response, and also around the minima around $4\,\mathrm{kHz}$ of the transfer function t_{h}, b_{h}.

The largest source of error between model and measurement appears to be the $50\,\mathrm{Hz}$ noise, causing up to 1 dB and 0.2 radians error. This error is still present despite the exclusion of several samples around $50\,\mathrm{Hz}$ and $100\,\mathrm{Hz}$. Further, noise is also present across the low frequencies from $1\,\mathrm{Hz}$ to $5\,\mathrm{kHz}$. Potentially with further averaging and improvement of the measurement setup – specifically the placement of the

equipment relative to mains power – a higher degree of accuracy could be achieved with the tone stack parameters.

Common-emitter amplifier

Following from the simulated combination analysis two sets of optimisations were performed, one using all signals and the other using only 1 & 3. For each combination the optimisation was repeated 30 times using initial parameter values taken from a uniform distribution of $\pm 10\%$ of the measured parameter values (and specified parameter values for the BJT), the results of the optimisation using all signals displayed in Table 7.5 and for the optimisation using signals 1 & 3 in Table 7.6.

Referring to both sets of results, a clear discrepancy can be observed between the results of the tone stack and those of the common-emitter amplifier: significantly more error is found for the common-emitter. This is likely due to the distortion in the output signal, the aliasing of the distortion, and the difficulty of the DAQ to accurately capture the distortion's high frequency content without phase error. Despite this increase in error, several linear parameters in both sets of results achieve a high degree of accuracy with the lowest being C_3 for both combinations.

Similar levels of error are observed in both sets of results, correlating with the results of the combination analysis. A notable difference between the two sets is that the results when using all signals provides BJT parameter values with an overall lower RSD than the results when only using sets 1 & 3. This aligns with the motivation of choosing the low amplitude signals (2 & 4), showing that including wideband signals with less distortion has improved the repeatability of estimating the BJT parameters.

One potential explanation for the generally higher error than the tone stack can be found in the higher than average RSD values for each BJT, for both combinations of signals. Most notably parameters I_s and β_r have significantly higher RSDs than the linear parameters suggesting that the objective function is insensitive to these parameters, and as such, it is possible that their inaccuracy has been compensated for by other parameters.

Displayed in Figure 7.15 is a comparison between the circuit measurements and the model using the selected $\boldsymbol{\theta}_\mathrm{ce}$ from the optimisation using all signals. For the low-amplitude signals, the resultant error is approximately 1 order of magnitude lower than that of the signal, though as the signal is of low amplitude the noise floor does not appear much lower. Further averaging may aid in the ability of the model to fit the circuit, but this would require more consideration as to the temperature changes of the BJT during the measurement to ensure the behaviour does not change.

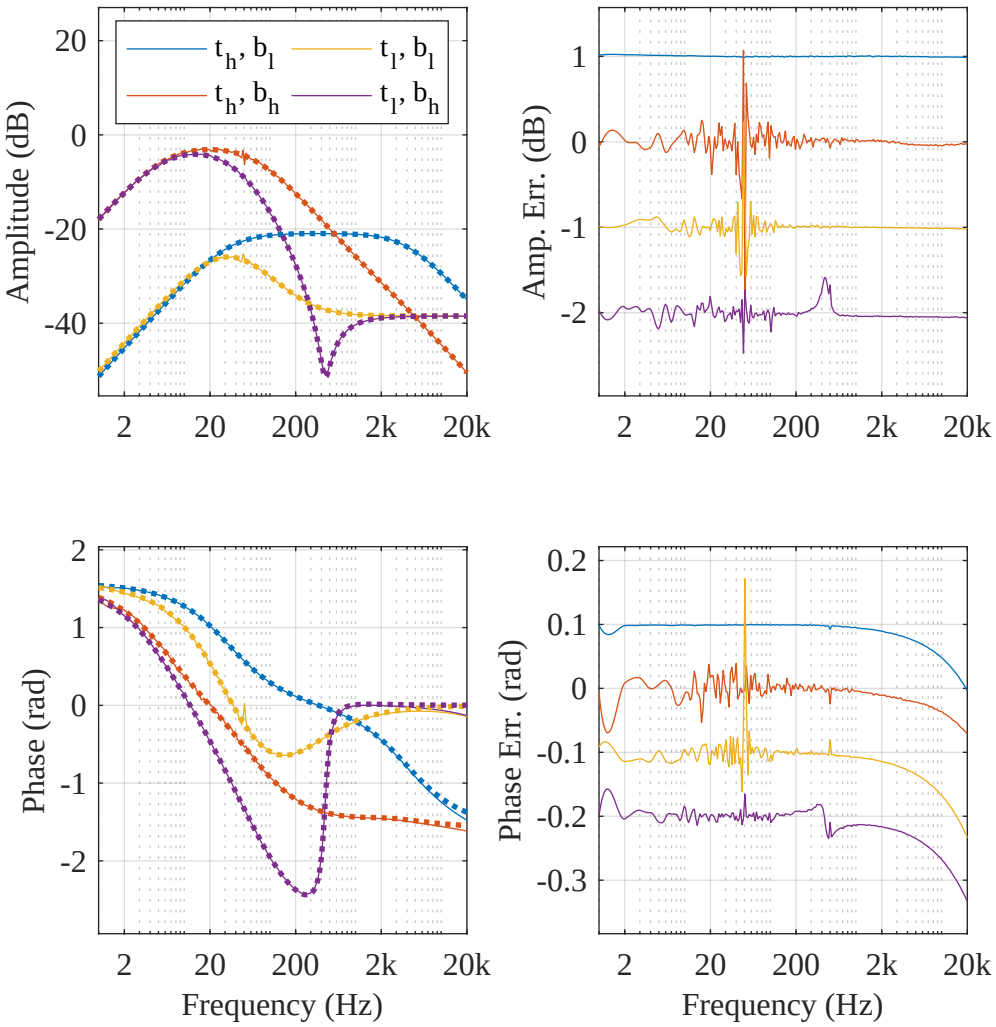


Figure 7.14: Optimised fit of the tone stack to measured data using values for $\boldsymbol{\theta}_{\text{ts}}$ as in Table 7.4. Error signals are offset by 1 dB and 0.1 rad to prevent overlap.

Table 7.5: Parameters of the common-emitter, from left to right: component symbol and units, the value specified by the schematic, the directly measured value using an LCR meter and its tolerance, the resulting θ_{ce} with the minimum ξ from the 30 sets, its error relative to the directly measured values, and the standard deviation of the full set. Resultant parameter values are from optimisations using all excitation signals.

$\boldsymbol{\theta}_{\text{ce}}$	Spec.	Direct	Tol.	Estim.	Error	RSD
R_1 (kΩ)	470	466.8	$\pm 1.06\%$	431.86	7.49%	0.599%
R_2 (kΩ)	68	68.18	$\pm 0.52\%$	70.868	3.94%	0.389%
R_3 (kΩ)	3.9	3.834	$\pm 0.35\%$	4.2550	11.0%	0.654%
R_4 (kΩ)	10	9.997	$\pm 0.32\%$	10.251	2.54%	0.496%
C_1 (nF)	4.7	4.555	$\pm 0.34\%$	4.6366	1.79%	0.391%
C_2 (nF)	1000	964.5	$\pm 0.32\%$	999.86	2.32%	0.296%
C_3 (nF)	10	10.114	$\pm 0.32\%$	10.226	1.14%	0.990%
I_{s} (pA)	10	-	-	2.7810	-	8.35%
N	1	-	-	1.0375	-	0.303%
β_{f}	200	-	-	190.98	-	1.64%
β_{r}	4	-	-	21.387	-	99.2%

Table 7.6: Parameters of the common-emitter as above. Resultant parameter values are from optimisations using excitation signals 1 & 3.

$\boldsymbol{\theta}_{\text{ce}}$	Spec.	Direct	Tol.	Estim.	Error	RSD
R_1 (kΩ)	470	466.8	$\pm 1.06\%$	462.14	0.999%	0.438%
R_2 (kΩ)	68	68.18	$\pm 0.52\%$	71.020	4.16%	0.270%
R_3 (kΩ)	3.9	3.834	$\pm 0.35\%$	4.3888	14.5%	2.57%
R_4 (kΩ)	10	9.997	$\pm 0.32\%$	11.298	13.0%	1.94%
C_1 (nF)	4.7	4.555	$\pm 0.34\%$	4.4917	1.39%	0.179%
C_2 (nF)	1000	964.5	$\pm 0.32\%$	865.00	10.3%	1.82%
C_3 (nF)	10	10.114	$\pm 0.32\%$	10.203	0.880%	1.71%
I_{s} (pA)	10	-	-	1.3559	-	1120%
N	1	-	-	1.0702	-	3.71%
β_{f}	200	-	-	161.57	-	2.17%
β_{r}	4	-	-	2.3059	-	498%

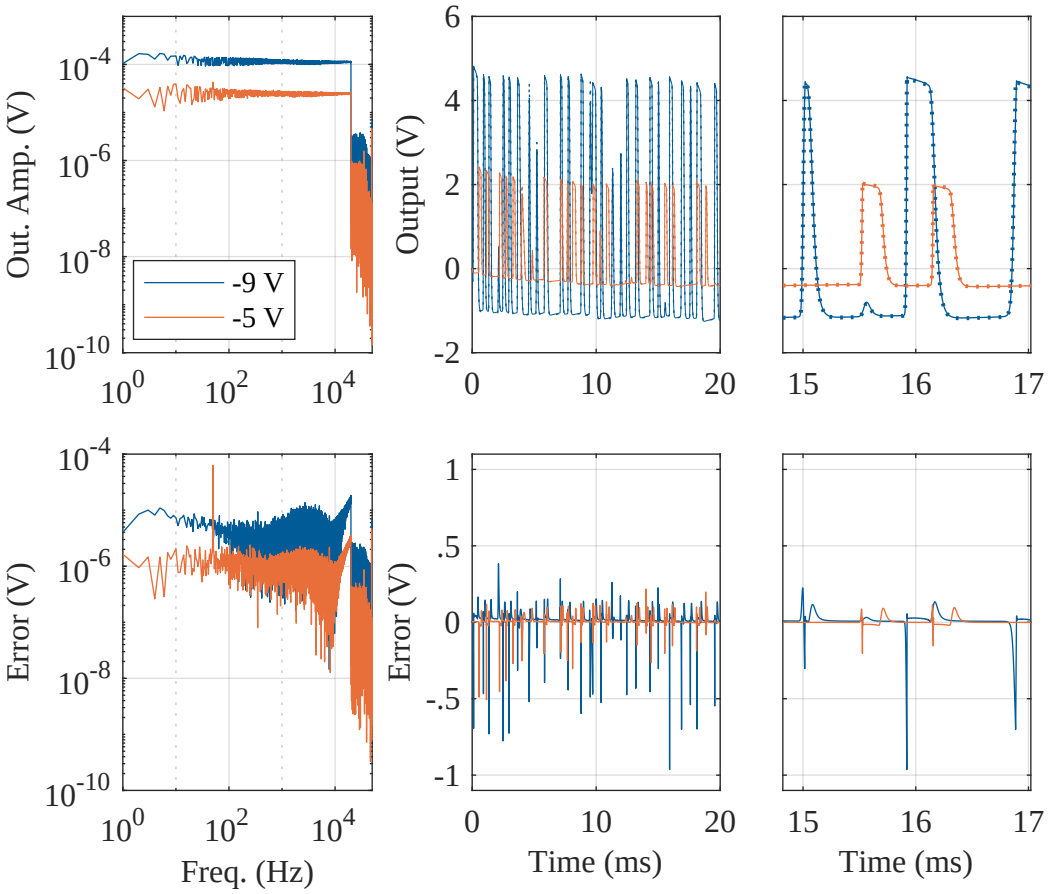


Figure 7.15: Optimised fit of the common-emitter amplifier to measured data using $\boldsymbol{\theta}_{\text{ce}}$ from Table 7.5. (left) Low amplitude signals and respective error, (middle) high amplitude signals and error, (right) zoom of the middle plot showing peak error.

High amplitude signals have a maximum error of approximately $1\,\text{V}_{\text{pp}}$. The far right plot focusses on the area in the signal at which this peak occurs, revealing that the majority of the error in the signal appears at its transitions. This result is further evidence of issues of matching the exact phase at high frequencies as was noted when modelling the DAQ.

7.3.3 Model validation through change of load

Inspecting the parameter error directly provides one perspective of the success of the parameter estimation, but a second perspective can be provided by validating the resultant models through changing the load of the circuit and model and observing the change in error. Should a model be able to adapt to a change in load without an increase in the output error then the estimated parameter values can be used in circuits

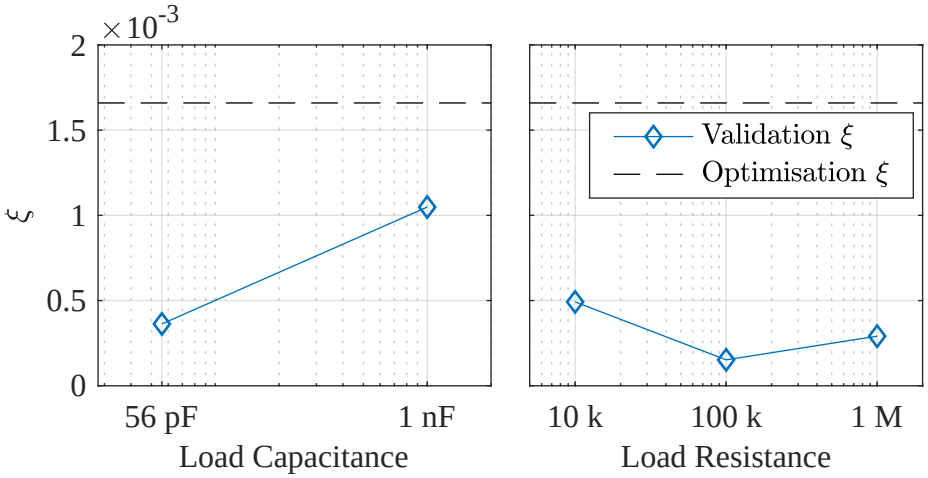


Figure 7.16: Validation of the tone stack by changing load component, comparing two changes in capacitance and three changes in resistance. The dashed line marks the optimised value of ξ with values of C_{ko} and R_{ko} as used for the optimisation.

that extend the chosen case study, for example, placing the tone stack in front of an amplifier circuit.

Tone stack

The tone stack was validated using 5 loads: 2 different capacitor values and 3 different resistors. These components were chosen to demonstrate some possible adaptations of the tone stack circuitry to provide different frequency responses, or the possibility of connecting a successive circuit. While the capacitance was changed the output resistance was omitted defaulting to R_{ai}, and while the resistance was changed the capacitance was omitted, defaulting to C_{ai}. All of the transfer functions of the tone stack were used in the validation of the model, with values in $\boldsymbol{\theta}_{\text{ts}}$ the same as select for Table 7.4.

The results of the validation are shown in Figure 7.16. The value of ξ is only ever lower than the resulting values of ξ from the optimisation. As such the validation is successful, suggesting that the estimated parameters are suitable for modified circuits/models.

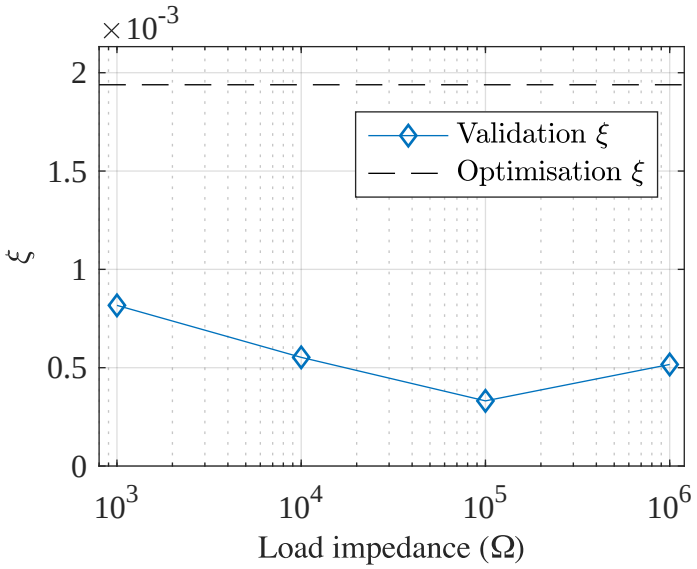


Figure 7.17: Validation of the common-emitter amplifier by changing load resistance. Four resistors were used to imitate the potential input impedances of a following guitar pedal. The dashed line marks the optimised value of ξ with values of C_{ko} and R_{ko} as used for the optimisation.

Common-emitter amplifier

Four different output resistors were used to validate the common-emitter amplifier, each selected at decades between $1\,\text{k}\Omega$ and $1\,\text{M}\Omega$, chosen to model the potential input impedances of a following guitar pedal. Capacitance was not changed from the optimisation. The set of $\boldsymbol{\theta}_{\text{ce}}$ chosen for the validation was from the results of the optimisation using all signals.

The values of ξ for each of the validation loads is beneath that of the optimised value of ξ, as displayed in Figure 7.17, suggesting that despite not achieving as high parameter accuracy as the tone stack, the retrieved values are still suitable for modelling modified circuits.

7.4 Conclusion

Identification of two circuits has been performed with the objective of estimation their physical circuit parameters. For the first, linear case – the tone stack – parameters were retrieved to a high degree of accuracy, which when used to model a circuit with a modified load resulted in no increase in error between model and measurement. For the second case – a common-emitter amplifier – less accuracy was achieved, likely

attributed to the nonlinear behaviour of the circuit. Resulting parameters were still applicable in the modelling of the amplifier with different loads, without increasing the error between model and measurement.

Simulated data was taken a step further than previously in the thesis, and used to predict the optimisation results of the measured data. Through the inclusion of AWGN general trends were captured for both circuits. This strategy could be repurposed to predict the number of averages required to achieve a noise floor that yields sufficiently low values of $\langle \epsilon_\theta \rangle$, given that no other non ideal effects present a limitation.

Two anomalous results in the prediction of $\langle \epsilon_\theta \rangle$ using simulated measurements was seen in the high amplitude signals for the common-emitter amplifier. A likely cause for this is the phase error in DAQ, further exacerbated by the high frequency transitions present in said signals. A direct solution to this would be further analysis and compensation of the DAQ to avoid this source of error. High frequency content in the common-emitter amplifier signals also caused issues with regards to aliasing. A preferable strategy would instead to be find a set of signals that reduces the need for high frequency content to achieve parameter retrieval, thus avoiding issues with aliasing and limitations of measurement equipment.

One shortcoming in the prediction of the results of the common-emitter amplifier is that there is no reference set of BJT parameters with which to compare. This forced the analysis to only use a subset of the parameters to predict the performance of the full set. A relative high amount of error was indicated in the results of the optimisations on measurements by the high values of RSD. For future study of the parameter estimation of audio circuits, a tool akin to an LCR meter for nonlinear devices is necessary.

Chapter 8

Conclusion

The main aim of this work has been to address the disparity between real guitar effect circuits and their corresponding physical circuit models. To that end two complementary identification strategies were proposed – calibration and parameter estimation – that utilise optimisation to reduce the error between the input/output measurements of a circuit and the simulated equivalent. The motivation was to provide modellers of guitar effects (and more broadly audio effects in general) with strategies to better relate their models to measurements of a real world equivalent. Results indicate that the proposed identification procedures can not only improve the fit of the model to the input/output behaviour but also be used to directly estimate the physical component parameters of the circuit.

8.1 Summary and contributions

Chapter 2 discusses the current capabilities of physical Virtual Analogue (VA) models and the parallel track of identifying circuits with black and grey box models, both progressing without intersection. Each electronic component present in case studies throughout the thesis are shown with their most commonly selected physical I-V relation, and it is demonstrated how by using MNA, circuits featuring these components can be converted into computable transfer functions and nonlinear state-space models. Existing investigations into the modelling of components for VA are discussed, noting the complexity of measuring devices, determining underlying behaviour, and extracting relevant physical parameters.

Chapter 3 details several root finding algorithms already used in solving the nonlinear equations present in physical VA models. These algorithms are analysed for their ability to converge within a time constraint, particularly in the multivariate case. The

initial objective stated for this work was to find algorithms which provide robust and efficient performance when simulating circuit models. Two extensions to Newton's method are introduced that use information derived from the form of the nonlinearity in state space models to improve the robustness of the algorithm. From those investigated, algorithms were determined that could successfully simulate both case studies over a range of input amplitudes. In the scenario that high-amplitude signals are driving the circuit models a demonstrable improvement was found in both the robustness and efficiency from the proposed algorithms. For the purpose of identification, should the excitation signal cause the circuit model to fail with traditional algorithms, the proposed algorithms provide an alternative which can extend the region of possible excitation signals.

Chapter 4 details an analysis of the germanium BJT, motivated by notably different behaviour when compared to a silicon equivalent. Analysis is performed through performing DC measurements upon germanium BJTs and extending the Ebers-Moll model until the curves are sufficiently fit. Resulting models of the OC44 and AC128 are placed in VA circuit models in which they were originally used, the Rangemaster and Fuzz Face. Several levels of complexity in the BJT model are compared to determine whether model extensions are required to properly model the BJT in physical VA models. Results indicate a marked difference between the Ebers-Moll model and the extended model. Should notable error be observed in a circuit model using BJTs, several options have been given with which to extend the BJT model to improve the fit to measurements of a real circuit.

Chapter 5 introduces the shared concepts behind the circuit identification strategies presented in the following chapters. The utilised data is restricted to simulated measurements to avoid noise and unmodelled behaviour that may appear in circuit measurements, placing the focus instead on the design and analysis of the identification problem. A vector of parameters is defined as the collection of each of the circuit's component parameters, the values of which are found through the application of an optimisation algorithm that changes the values to minimise the difference between input/output measurements of a circuit and its model. Parameter redundancy is detected in the input/output circuit models. How to handle this redundancy is different for the two identification strategies: for calibration, a parameter is fixed to reduce the computational cost of the optimisation, selected by finding the parameter to which the objective function is least sensitive. For parameter estimation, an additional component is added with known parameter values which alleviates the redundancy, producing demonstrably retrievable parameter values through optimisation. Designing and

testing an excitation signal, objective function, and suitable models for optimisation facilitates the following work in calibration and parameter estimation.

Chapter 6 completes the study of calibration within the thesis through application to measurements of a Dallas Rangemaster Treble Booster circuit. Calibration is shown to dramatically improve the capture of the input/output behaviour of the circuit with the model in comparison to nominal parameters from a schematic. Validation metrics show values of ξ that are up to $20\times$ lower for the optimised model when processing sinusoids over a range of frequency/amplitude combinations. Analysing the parameters to which the objective function is least sensitive enables an optimisation time over $2.5\times$ faster for the Rangemaster model by fixing the value of 5 parameters, although also causing a marginal increase in error between circuit and model. Secondary to the calibration of the Rangemaster is a comparison between germanium and silicon BJTs within the circuit, with results indicating that the circuit model is better at fitting the behaviour of the Rangemaster with the germanium BJT than that using the silicon BJT. Overall the calibration strategy is shown to successfully reduce the error between the input/output relationship of a circuit and its model.

Finally, Chapter 7 applies the parameter estimation strategy to two real case study circuits, the first a tone stack featuring multiple audio parameters and the second the common-emitter amplifier with nonlinear behaviour. Parameter values are retrieved to within 2.5% of their directly measured value for the tone stack, with most of the parameter values falling within the tolerance of the LCR meter used to directly measure each component. Moderately accurate parameter values for the common-emitter amplifier are found, limited by aliasing and noise. Both cases exhibit no increase in error when modelling modified circuits demonstrating that the estimated parameter values could be used when combining the circuits with additional circuitry, for example in a full amplifier or guitar-pedal signal chain. The identification process is shown to be capable of estimating the component parameters of circuits though further work is necessary to provide similar accuracy for the nonlinear common-emitter amplifier as with the linear tone stack.

8.2 Future work

Due to their essential role in physical circuit models, the development of iterative solvers will remain an important topic of research. Specifically for identification, there may be further necessity of particularly robust algorithms. Consider the case that a circuit is highly nonlinear, and that to properly expose the behaviour of said circuit re-

quires driving the circuit with high-amplitude, high-frequency signals. Simulation failure may occur for models with certain parameter values, preventing the optimisation algorithm from exploring a region of the search space that may be where the optimal parameter set lies. This possibility has not arisen within the case studies presented in this work, but may yet be encountered in those used in future work. Additional consideration would then be necessary should a method like the New Iterate be selected, as it was found that for the Fuzz-Face circuit that the new initial iterate often caused non-convergence.

Vintage circuits feature a vast quantity of unique quirks that form the sonic palette of musicians, many of which require individual attention. In this work, this has been the germanium BJT, researched for its usage in the Rangemaster. Further examples of this are, to name a few: germanium diodes, nonlinear behaviour of inductors, for example, in the Cry Baby wah pedal, and specific op-amps with limited slew rate as used in the Proco RAT. An argument could be made for an archival investigation into such devices – time will reduce the number of functioning devices, and without digitising through accurate models, there may be a future in which none can recreate the guitar tones of the music of yesteryear.

As the field progresses larger circuits will become possible to simulate in real time, facilitated by the increase in available computational power. The next stages of research into the calibration strategy presented in this work should be in its application to circuits with more complexity and higher numbers of parameters. Presence of local minima may become a larger issue with the increase in search space dimensions. Application of parameter screening may offer a solution, not only reducing the cost of the optimisation algorithm, but reducing the complexity of the search space by minimising the number of dimensions it spans.

A wealth of future work is possible around the subject of parameter estimation for nonlinear circuits. A limit has been found in this work: the accuracy of the parameters retrieved for the common-emitter amplifier was significantly lower than that of the tone stack, making overcoming the issues caused by nonlinearities an interesting challenge. Investigation of the parameter estimation strategy is limited by the expense of the model simulation which is repeated thousands of times in the optimisation algorithm. Model complexity is a trade off between the range over which the model accurately captures behaviour and computation time. One solution to the issue of model complexity is referred to as ‘space mapping’ in which a ‘coarse’ model that is less computationally demanding is used with optimisation to find good estimates of the parameters, before transitioning to a ‘fine’ model which can be used to achieve higher

accuracy estimates [112]. For the common-emitter amplifier this could be found in the utilisation of a linear BJT model so that the circuit can be modelled as a transfer function, then moving to the complete state space model for details about the nonlinearity.

Both noise and distortion limited the success of estimating the parameters of the common-emitter amplifier. The inherent problem is that both signal to noise ratio present in the circuit measurement and aliasing from harmonic distortion in the output of the circuit model typically increase relative to the input signal amplitude. An ideal measurement strategy would find the sweet-spot at which aliasing is minimised and signal to noise ratio is maximised, or utilises a divide-and-conquer approach. The solution may be as simple as using more specifically designed hardware, with a lower noise range over the audio band, though this solution adds an additional hurdle to those wanting to utilise the identification process. An alternative is to use anti-aliasing strategies implemented directly in the physical circuit model, e.g. [56].

Different values of audio parameters have been used in a single model of the tone stack when estimating its parameters. By including multiple audio parameter values in the identification – should the process be successful – it is given that the model accurately represents the circuit at these values. What is not certain is how accurately the model fits to measurements of a circuit at values in between the audio parameter values used in the measurements. Potentiometers specifically have different laws as to how resistance changes as the wiper passes over the track. An interesting study would be to identify these potentiometer laws such that the model can accurately represent the circuit at any given combination of audio parameter values.

Appendix A

OC44 Datasheet

R.F. JUNCTION TRANSISTOR OC44

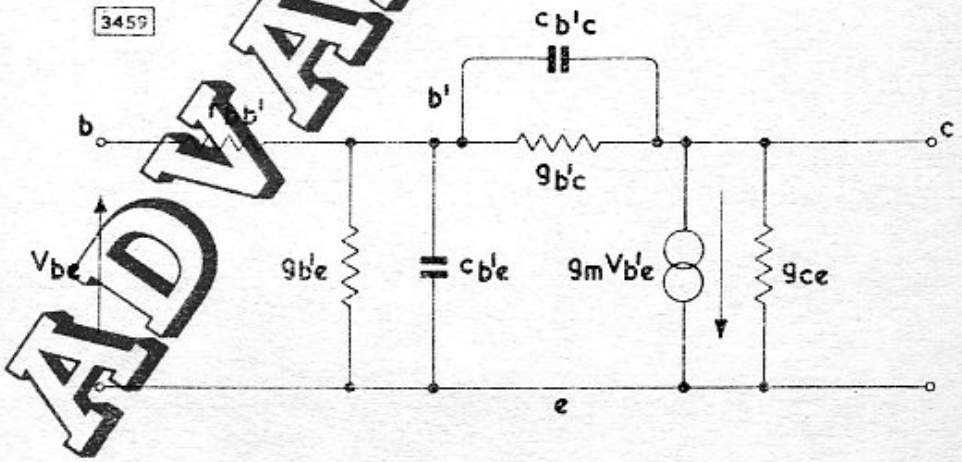
R.F. junction transistor of the p-n-p alloy type in all-glass construction intended for use in converters and mixer-oscillator circuits.

TYPICAL CHARACTERISTICS (measured at Tambient = 25°C)

Grounded Base		Min.	Av.	Max.	
Collector leakage current (Vc=-2V)	Ic(o)		0.5		µA
Collector leakage current (Vc=-15V)	Ic(o)			10	µA
Emitter leakage current (Ve=-2V)	Ie(o)		0.4		µA
Emitter leakage current (Ve=-12V)	Ie(o)			40	µA
Current amplification cut-off frequency (Measured at Vc=-6V, Ic=1mA)	fα	7	15		Mc/s
Grounded Emitter					
Collector leakage current (Vc=-2V)	I'c(o)		2		µA
Current amplification factor (Vc=-6V, Ic=1mA, f=1kc/s)	α'		100		

SMALL SIGNAL CHARACTERISTICS

Measured at: Vc		-6	V
Ic		1	mA
Equivalent circuit parameters (hybrid π network)			
Collector-to-base capacitance	$c_{b'c}$	10.5	pF
Base-to-emitter capacitance	$c_{b'e}$	410	pF
Collector-to-emitter conductance	g_{ce}	40	µmhos
Collector-to-base conductance	$g_{b'c}$	<0.5	µmhos
Base-to-emitter conductance	$g_{b'e}$	390	µmhos
Internal base resistance	$r_{bb'}$	110	Ω
Intrinsic transconductance	g_m	39	mA/V



b' indicates the internal base connection

ISSUE 1
12.56

Issued by **MULLARD LIMITED**
Data and Publications Section, T.S.D.

LIMITING VALUES (absolute ratings)

The equipment designer must ensure that no transistor exceeds these ratings and in arriving at the actual operating conditions, variations in supply voltages, component tolerances and ambient temperature must also be taken into account.

Grounded Base

Vc max.	-10	V
vc(pk) max.	-15	V

Grounded Emitter

*Vc max.	-10	V
*vc(pk) max.	-15	V

* These values apply with an external base to emitter resistance of <1kΩ. For other values of resistance, see curve on page 7.

Reverse Emitter Voltage

Ve max.	-8	V
ve(pk) max.	-12	V

Collector Current

Ic max.	5	mA
ic(pk) max.	10	mA

Emitter Current

Ie max.	5	mA
ie(pk) max.	10	mA

Collector Dissipation

pc max.	20	mW

TEMPERATURE RATINGS

Storage temperature	-55 to 70	°C
Max. junction temperature	65	°C
Junction temperature rise above ambient in free air	0.5	°C/mW

OPERATING NOTES

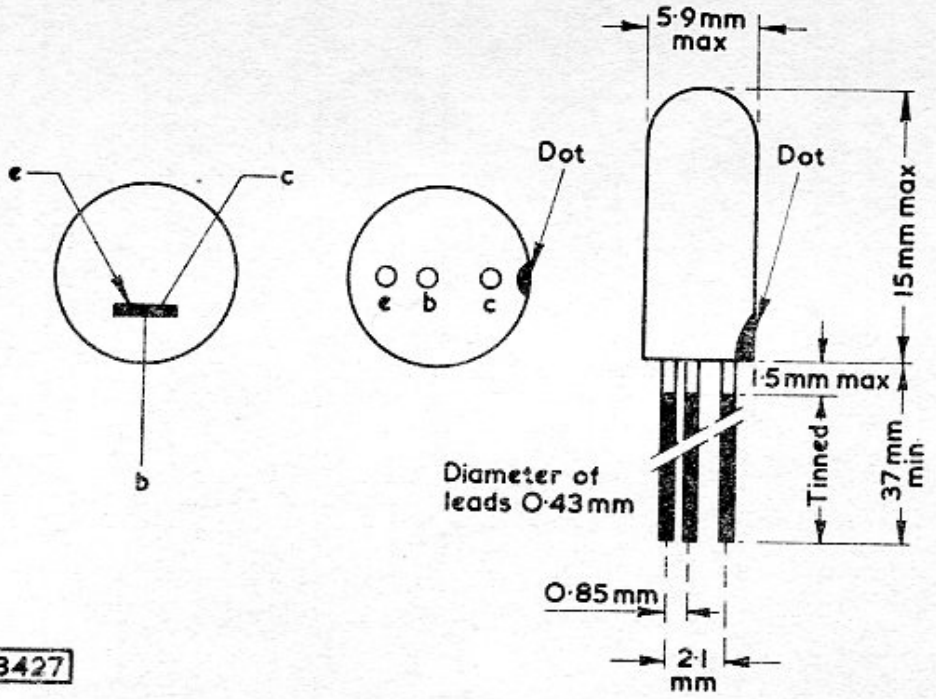
1. Care should be taken not to bend the leads nearer than 1.5mm to the seal.

2. Transistors are inherently sensitive to incident illumination; care should be taken to ensure that the external coating is not damaged.

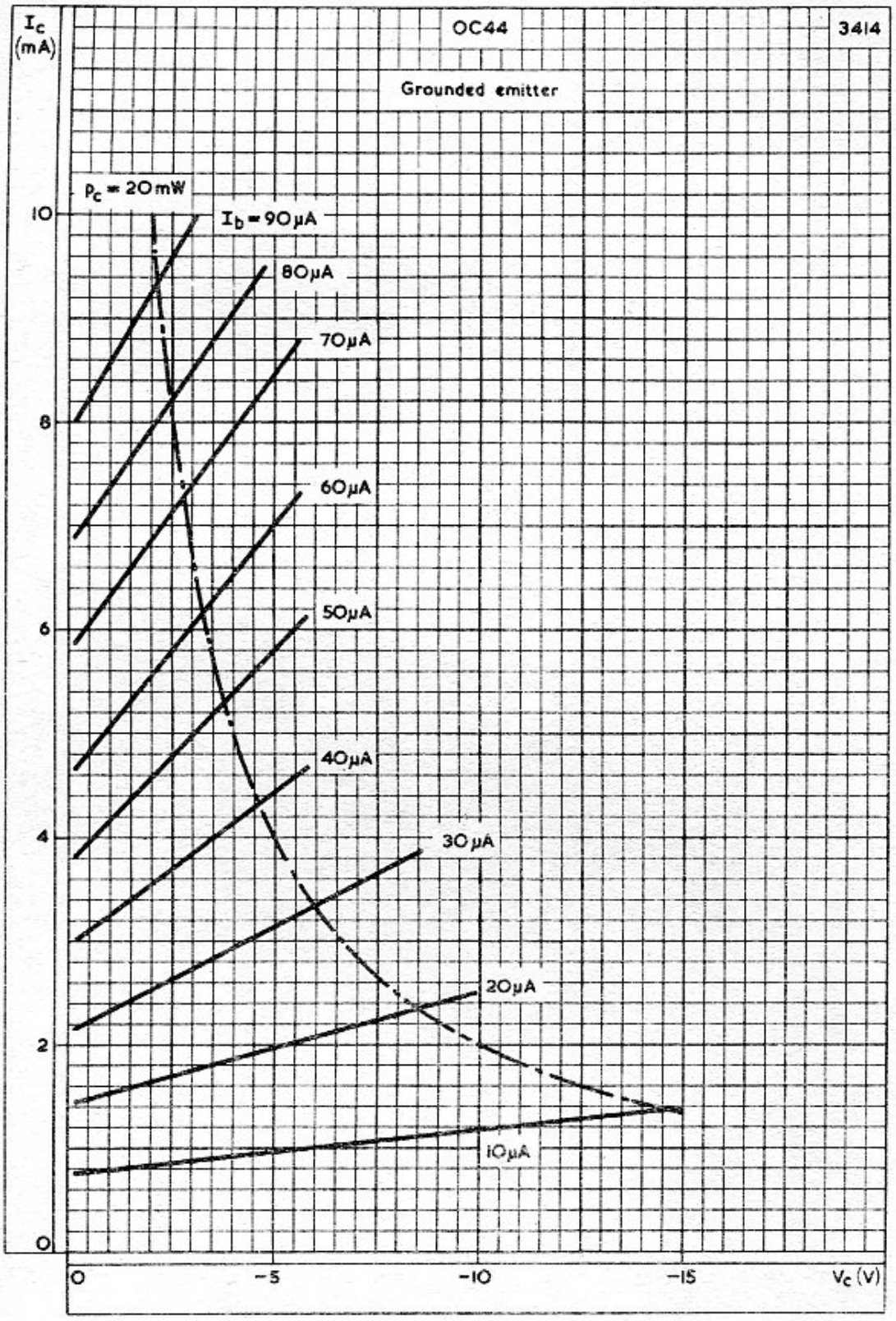
Mullard

Issued by MULLARD LIMITED
Data and Publications Section, T.S.D.

R.F. JUNCTION TRANSISTOR OC44

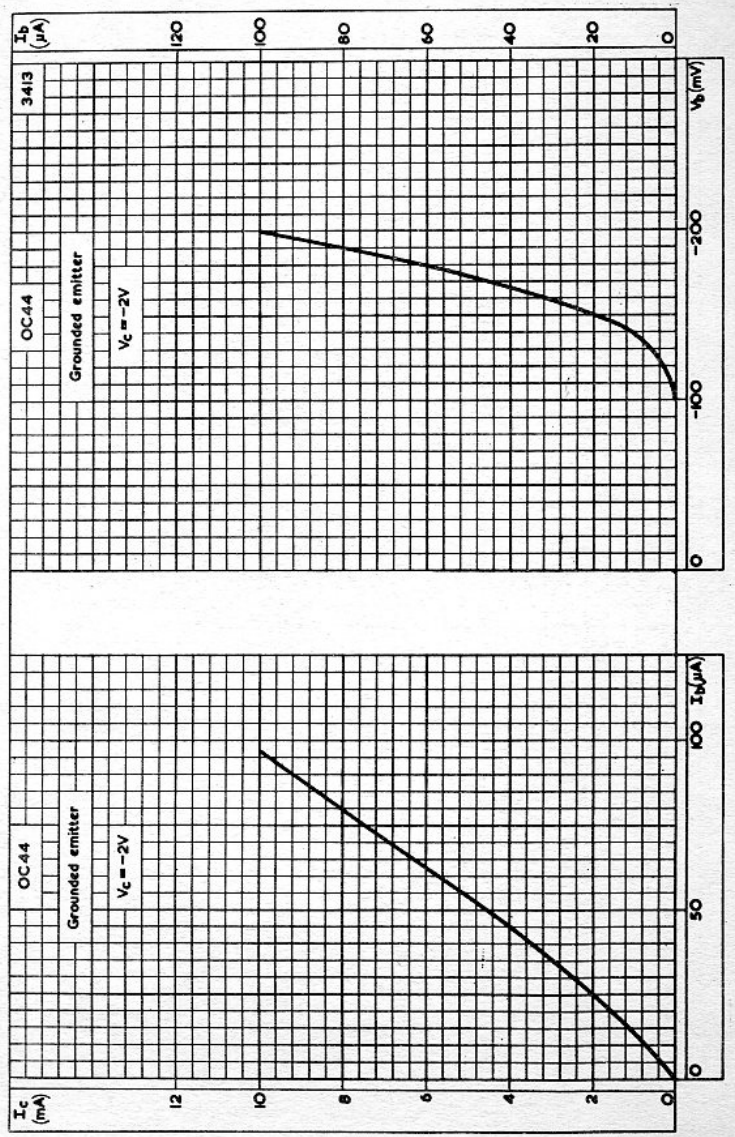


Issued by **MULLARD LIMITED**
Data and Publications Section, T.S.D.

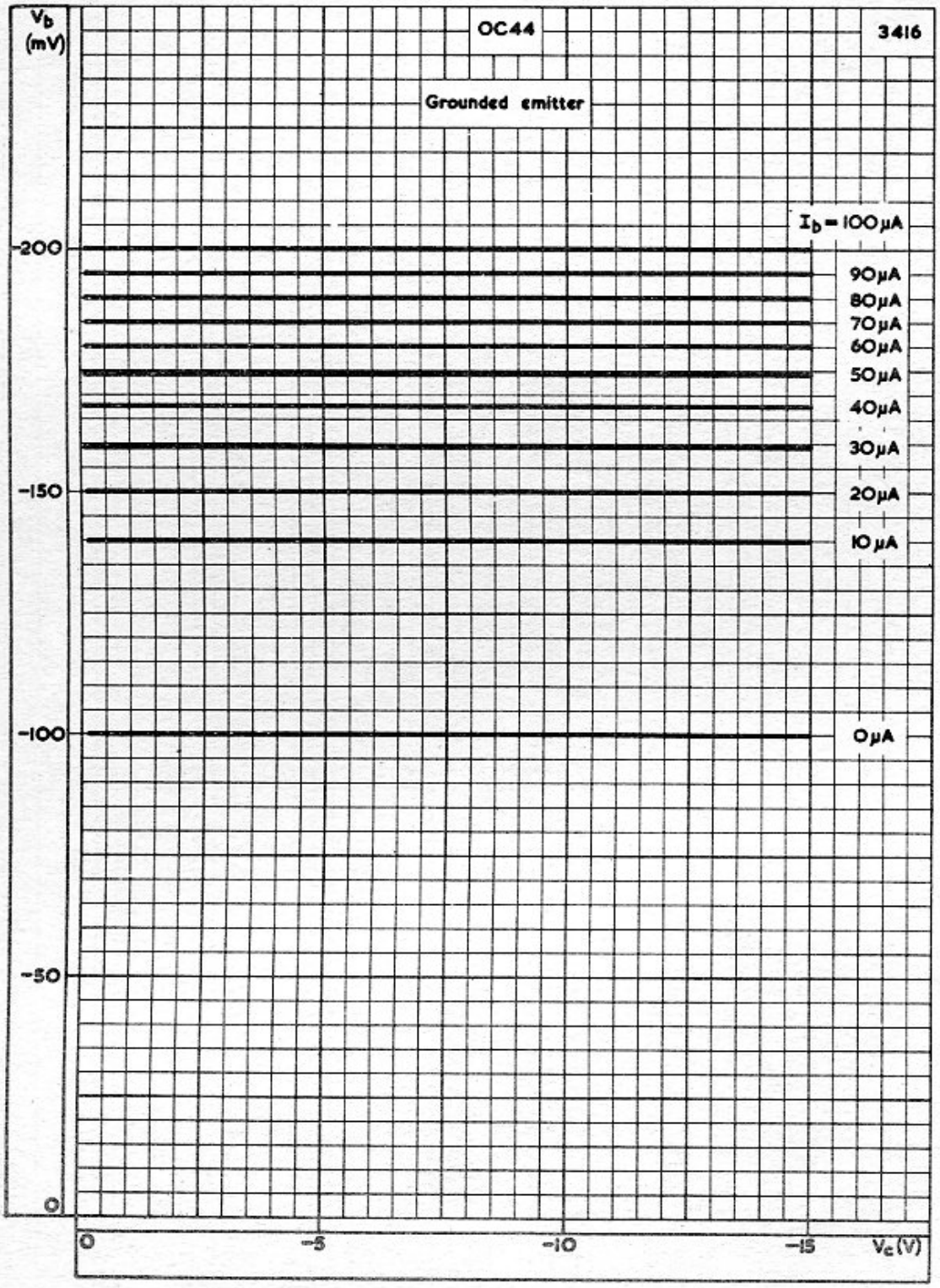


Mullard

Issued by **MULLARD LIMITED**
Data and Publications Section, T.S.D.

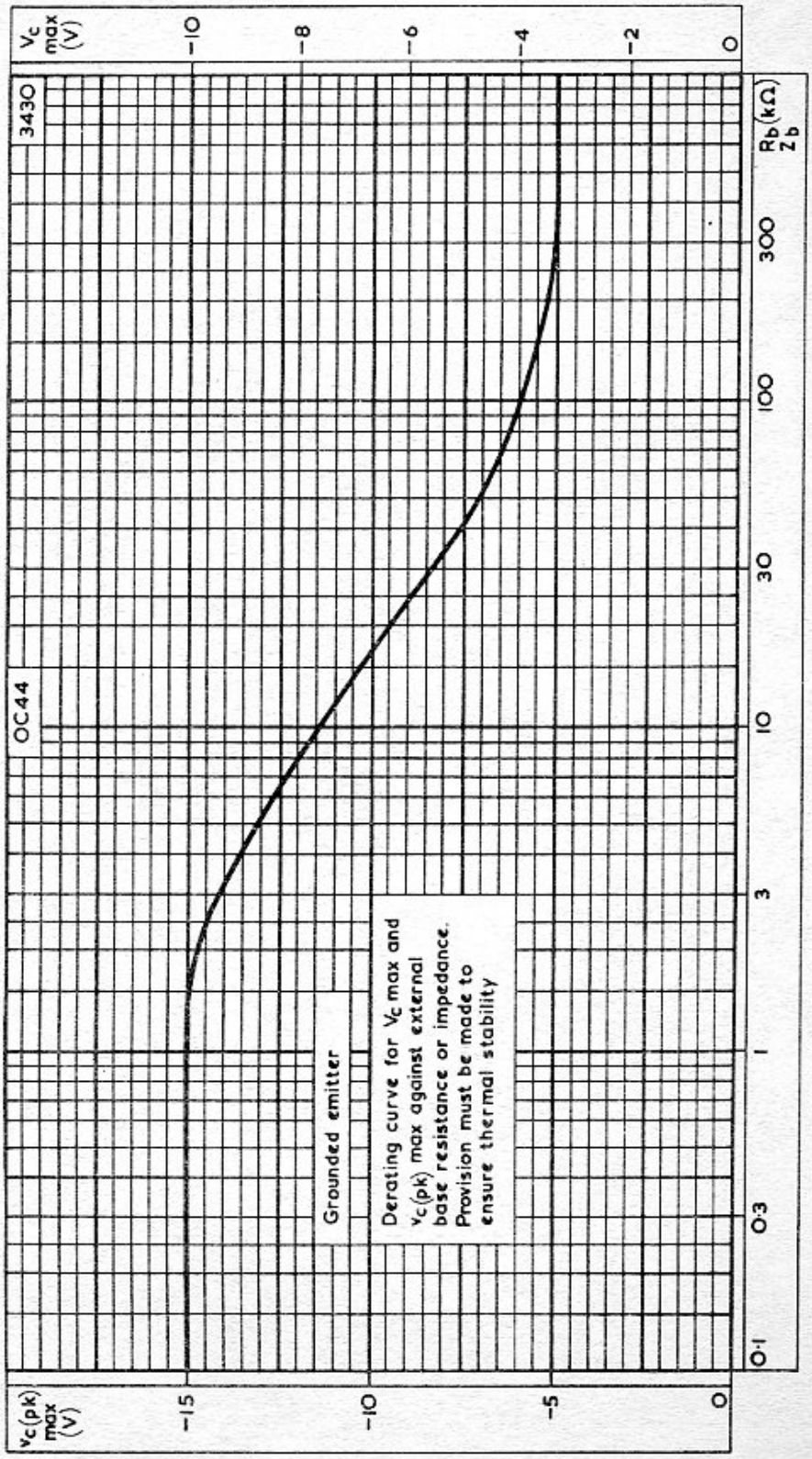


Issued by **MULLARD LIMITED**
Data and Publications Section, T.S.D.



Mullard

Issued by **MULLARD LIMITED**
Data and Publications Section, T.S.D.



Issued by **MULLARD LIMITED**
Data and Publications Section, T.S.D.

Appendix B

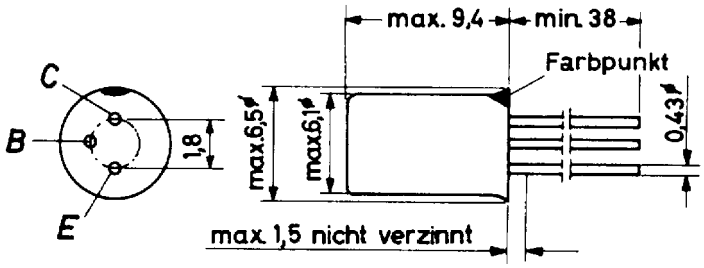
AC128 Datasheet

AC 128

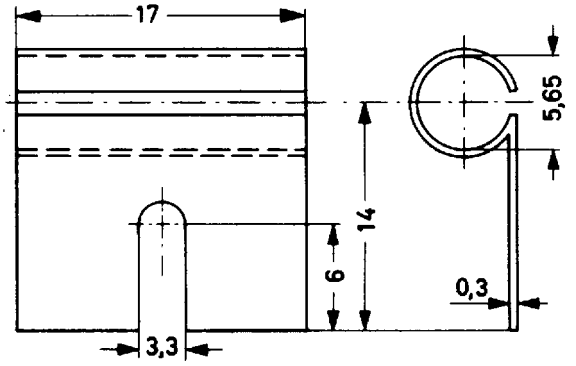
GERMANIUM - p-n-p - NF-TRANSISTOR
für Endstufen,
als Transistorpaar für Gegentakt-B-Schaltungen,
in Verbindung mit AC 127 als komplementäres Paar

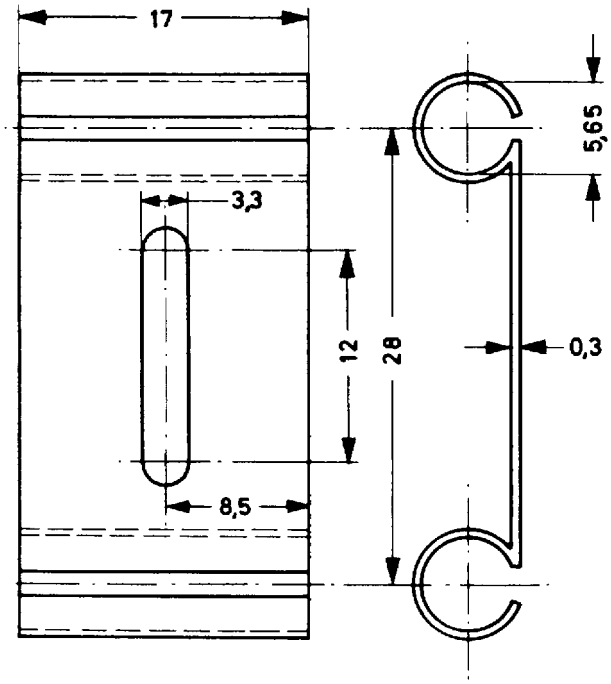
Abmessungen in mm:

Gehäuse: Metall
Roter Punkt: Kollektorseite



Kühlschelle 56 227





Kühlschelle 56 226

Wärmewiderstand:

K_G	$\leqq$	0,04 grd/mW	
K	$\leqq$	0,29 grd/mW	ohne Kühlschelle
K	$\leqq$	0,14 grd/mW	mit Kühlschelle
K	$\leqq$	0,08 grd/mW	mit Kühlschelle und Kühlfläche von 12,5 cm^2

Absolute Grenzwerte:

$-U_{CB}$	= max.	32	V	
$-U_{CE}$	= max.	32	V	1)
$-U_{EB}$	= max.	10	V	
$-I_C$	= max.	1	A	
$-I_B$	= max.	40	mA	
P	= max.	700	mW	2)
ϑ_j	= max.	90	°C	
ϑ_s	= min.	-55	°C	
ϑ_s	= max.	75	°C	

1) bei $R_{BE} \leqq 500\ \Omega$, siehe auch Grenzkurve

2) gesamte Verlustleistung; $\vartheta_{j\ max}$ darf dabei nicht überschritten werden. Spitzenwert bei B-Verstärkern mit Sprache- und Musik-Aussteuerung max. 1 W.

AC 128

Kennwerte: (ϑ_j = 25 °C, sofern nicht anders angegeben)

Kollektor-Reststrom bei $-U_{CB}$ = 10 V:	$-I_{CB\ 0}$	≦	10	µA
Emitter-Reststrom bei $-U_{EB}$ = 5 V, ϑ_j = 75°C:	$-I_{EB\ 0}$	≦	500	µA
Kollektorspannung bei $-I_{CB\ 0}$ = 200 µA:	$-U_{CB}$	≧	32	V
Emitterspannung bei $-I_{EB\ 0}$ = 200 µA:	$-U_{EB}$	≧	10	V
Kollektor-Restspannung bei $-I_C$ = 1 A [1]:	$-U_{CE\ 0}$	≦	0,6	V
Basisspannung bei U_{CB} = 0, I_E = 50 mA:	$-U_{BE}$	≦	300	mV
bei U_{CB} = 0, I_E = 300 mA:	$-U_{BE}$	≦	450	mV

Gleichstromverstärkung

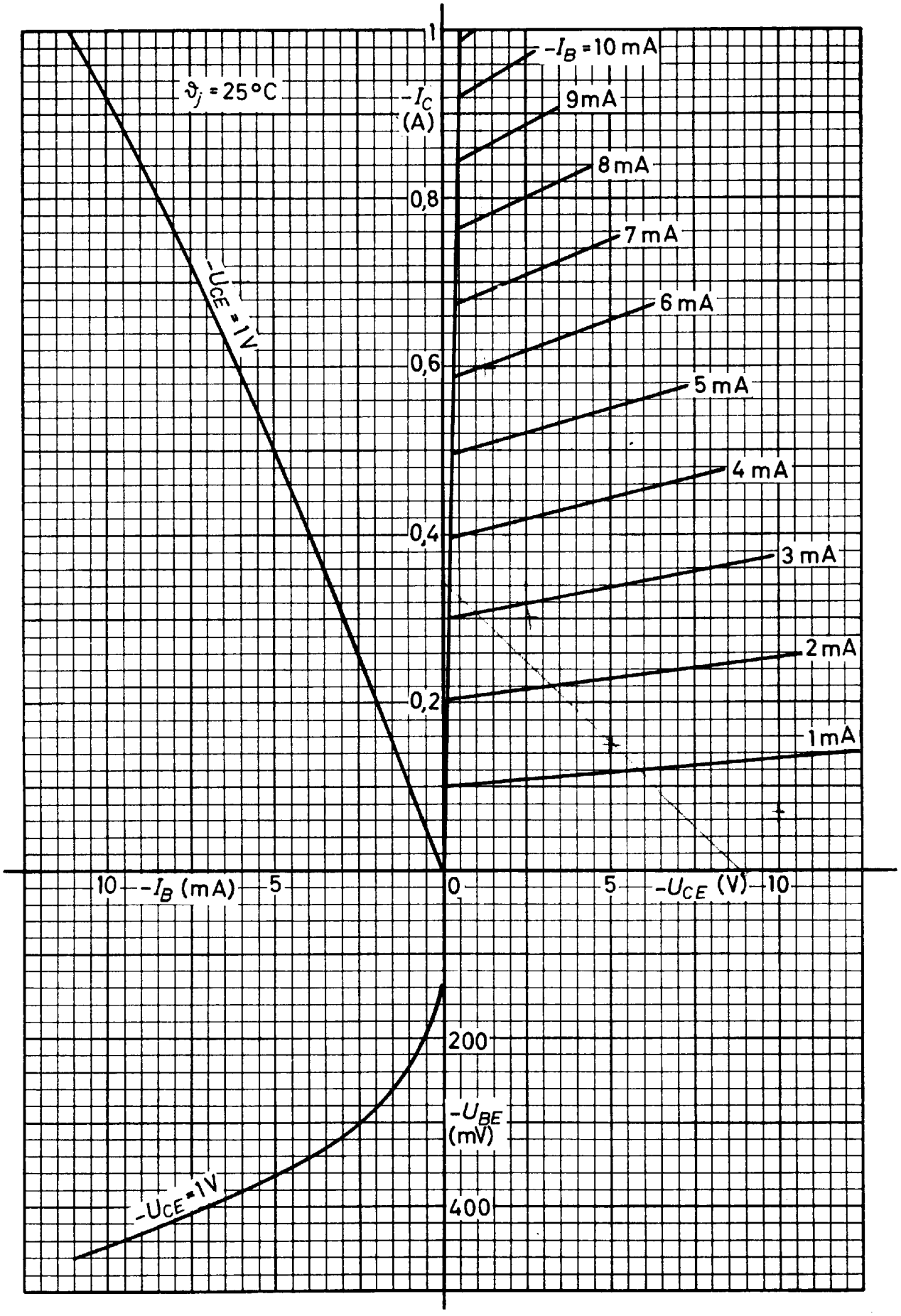
bei U_{CB} = 0, I_E = 50 mA:	B	= 90 (55...175)
bei U_{CB} = 0, I_E = 300 mA:	B	= 90 (60...175)
bei U_{CB} = 0, I_E = 1 A:	B	= 80 (45...165)
Frequenz für $\|\beta\|$ = 1 bei $-U_{CB}$ = 2 V, I_E = 10 mA:	f_1	= 1,5 (≧ 1,0) MHz
Grenzfrequenz bei $-U_{CB}$ = 2 V, I_E = 10 mA:	f_β	= 15 (≧ 10) kHz
Basisbahnwiderstand bei $-U_{CB}$ = 5 V, I_E = 1 mA:	$r_{bb'}$	= 25 Ω
Kollektorkapazität bei $-U_{CB}$ = 5 V, I_E = 0:	$C_{b'c}$	= 100 pF
Stromverstärkungs-Verhältnis $\frac{v_i\ (-I_C=500mA)}{v_{i\ max}}$		= 0,60 (≧ 0,50) [2]

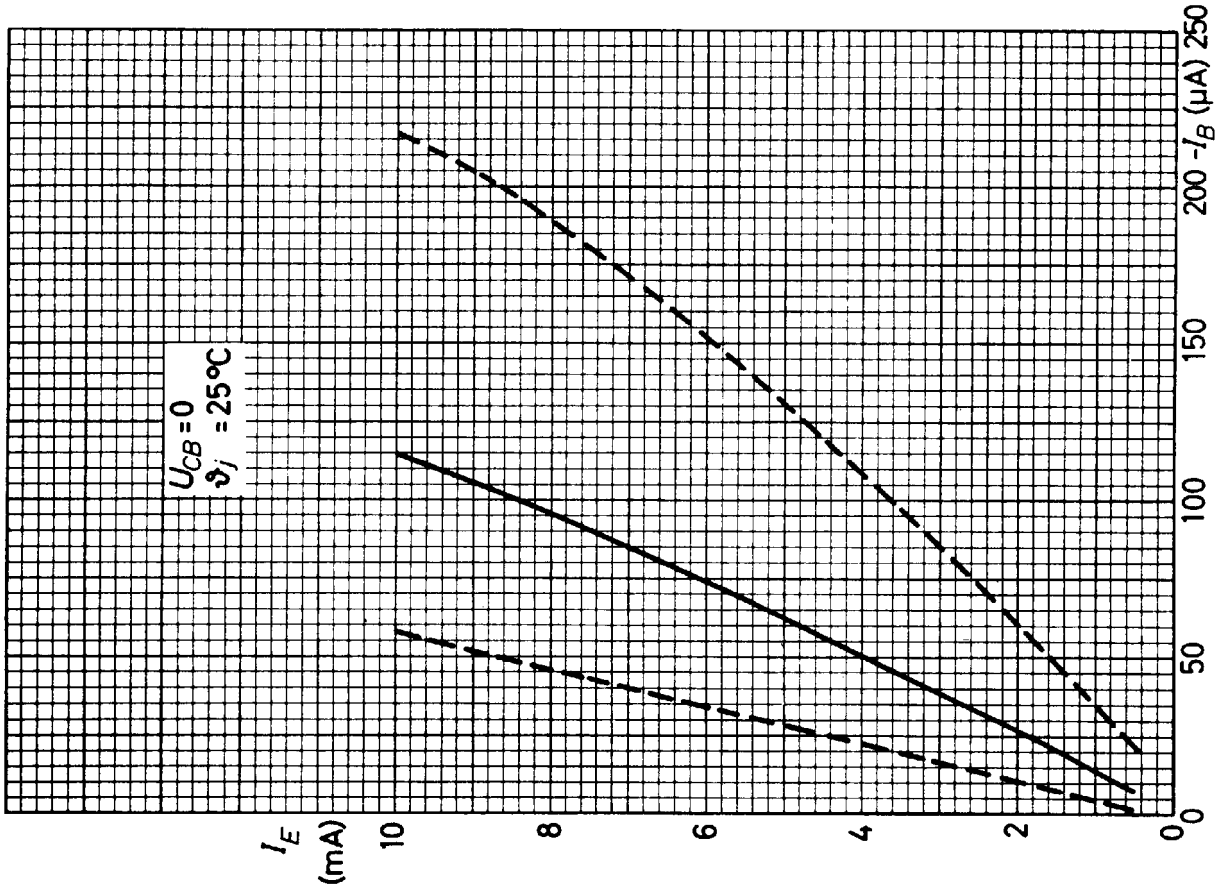
Transistorpaar:
Das Verhältnis der Gleichstromverstärkungen beider Transistoren zueinander bei I_E = 50 mA, U_{CB} = 0 sowie bei I_E = 300 mA, U_{CB} = 0 ist 1,1 (≦ 1,25).

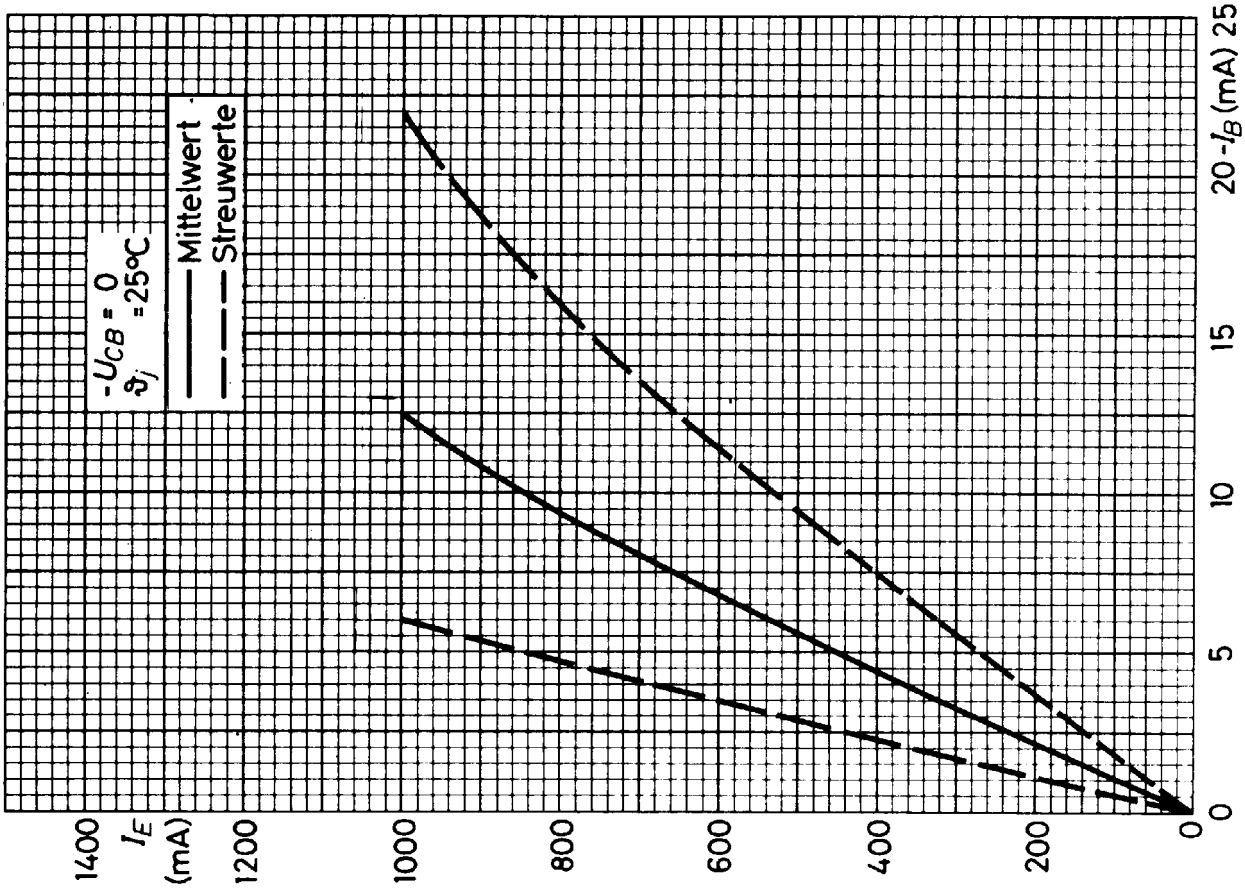
---

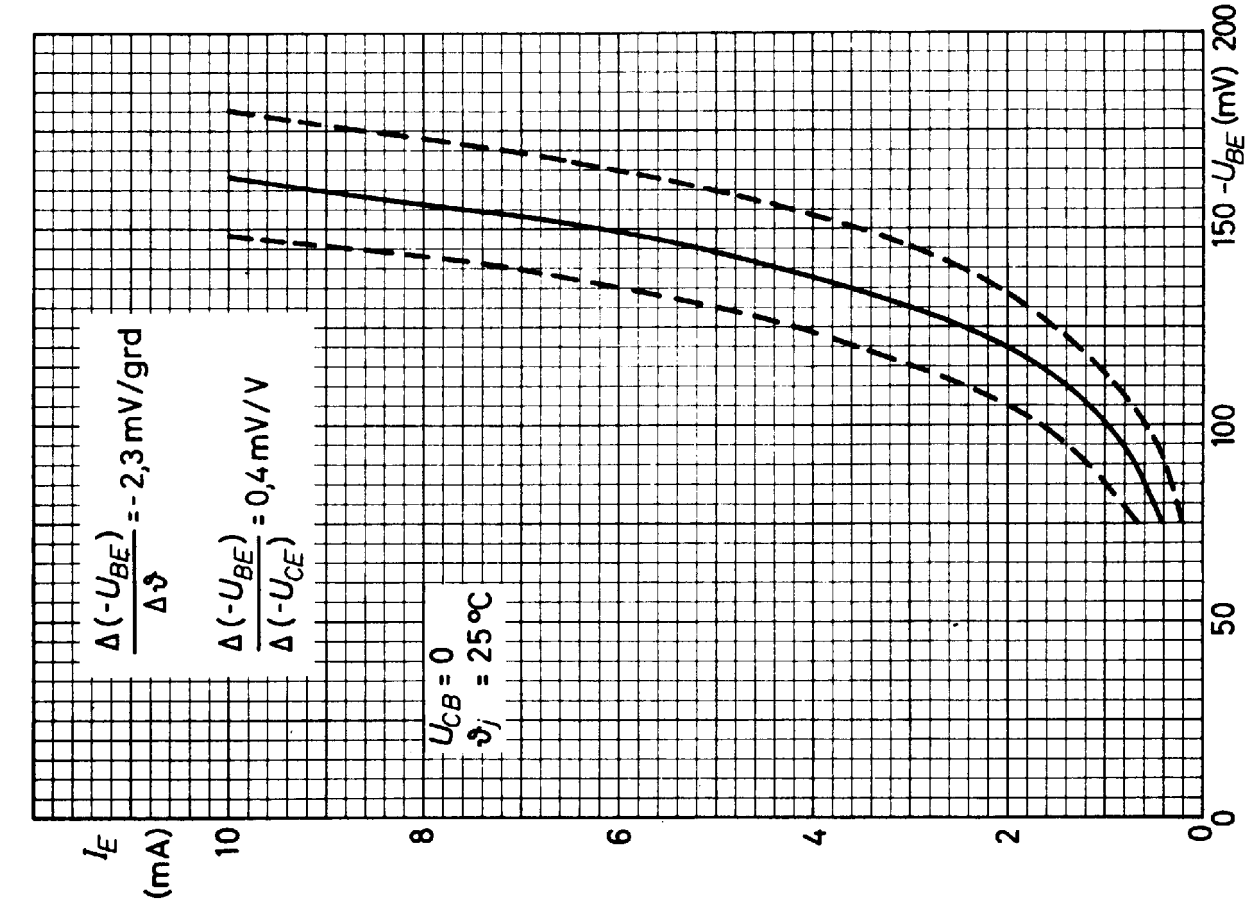
[1]) für die Kennlinie, die bei gleichem Basisstrom durch den Kennlinienpunkt $-I_C$ = 1,1 A, $-U_{CE}$ = 1 V geht

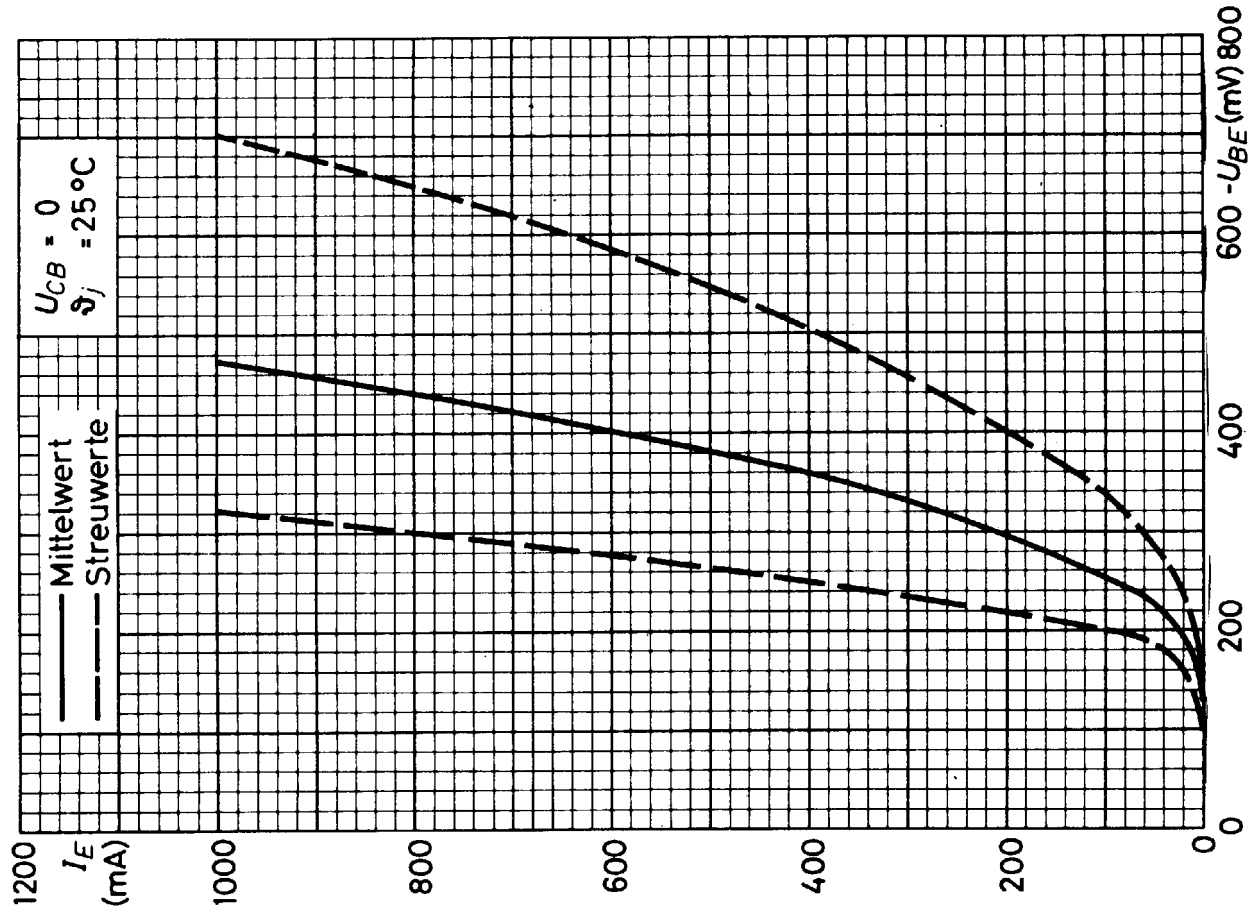
[2]) bei U_0 = 10 V, R_L = 16 Ω

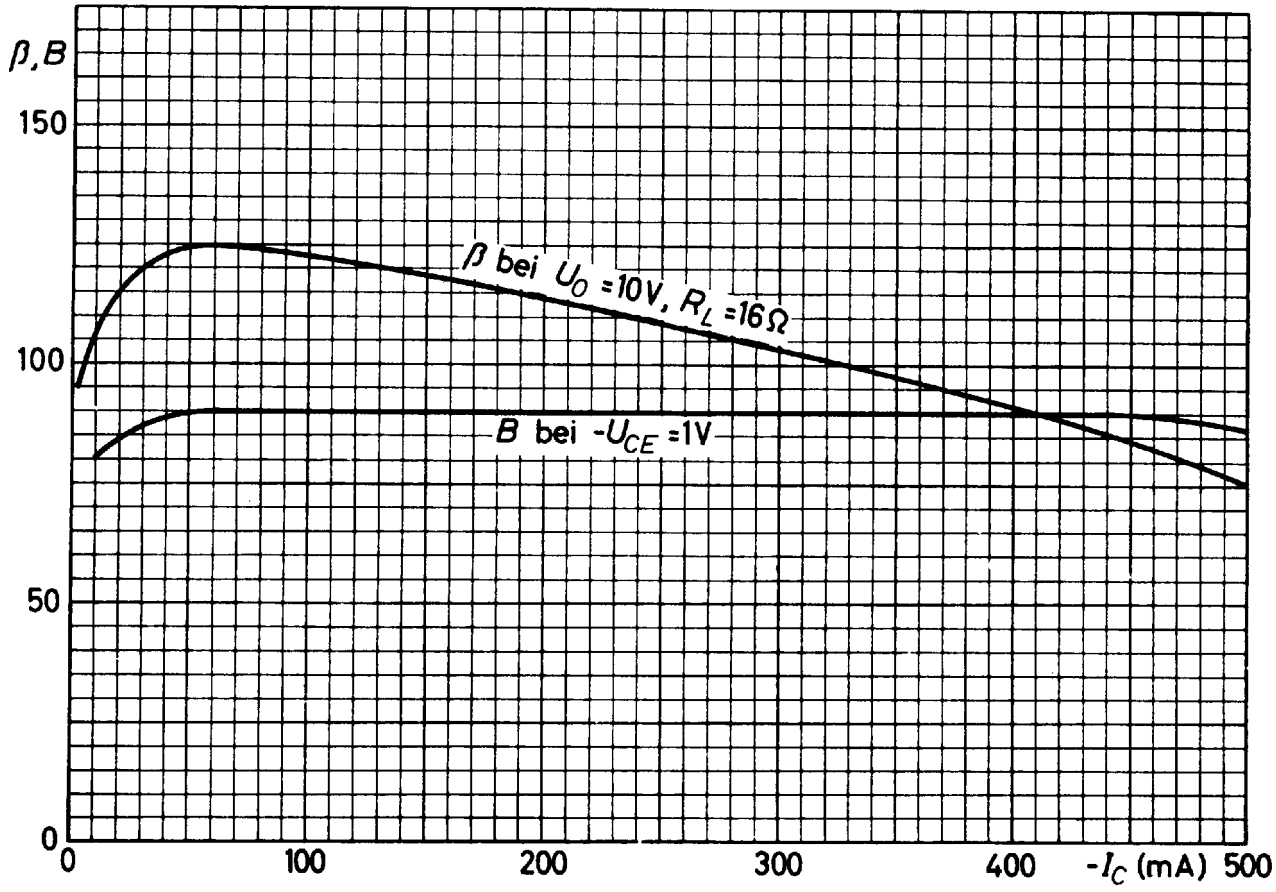

ϑ_j = 25 °C
$-I_B$ = 10 mA
9 mA
8 mA
7 mA
6 mA
5 mA
4 mA
3 mA
2 mA
1 mA
$-I_C$ (A)
1
0,8
0,6
0,4
0,2
$-U_{CE}$ = 1 V
10
$-I_B$ (mA)
5
0
5
$-U_{CE}$ (V)
10
200
$-U_{BE}$ (mV)
400
$-U_{CE}$ = 1 V

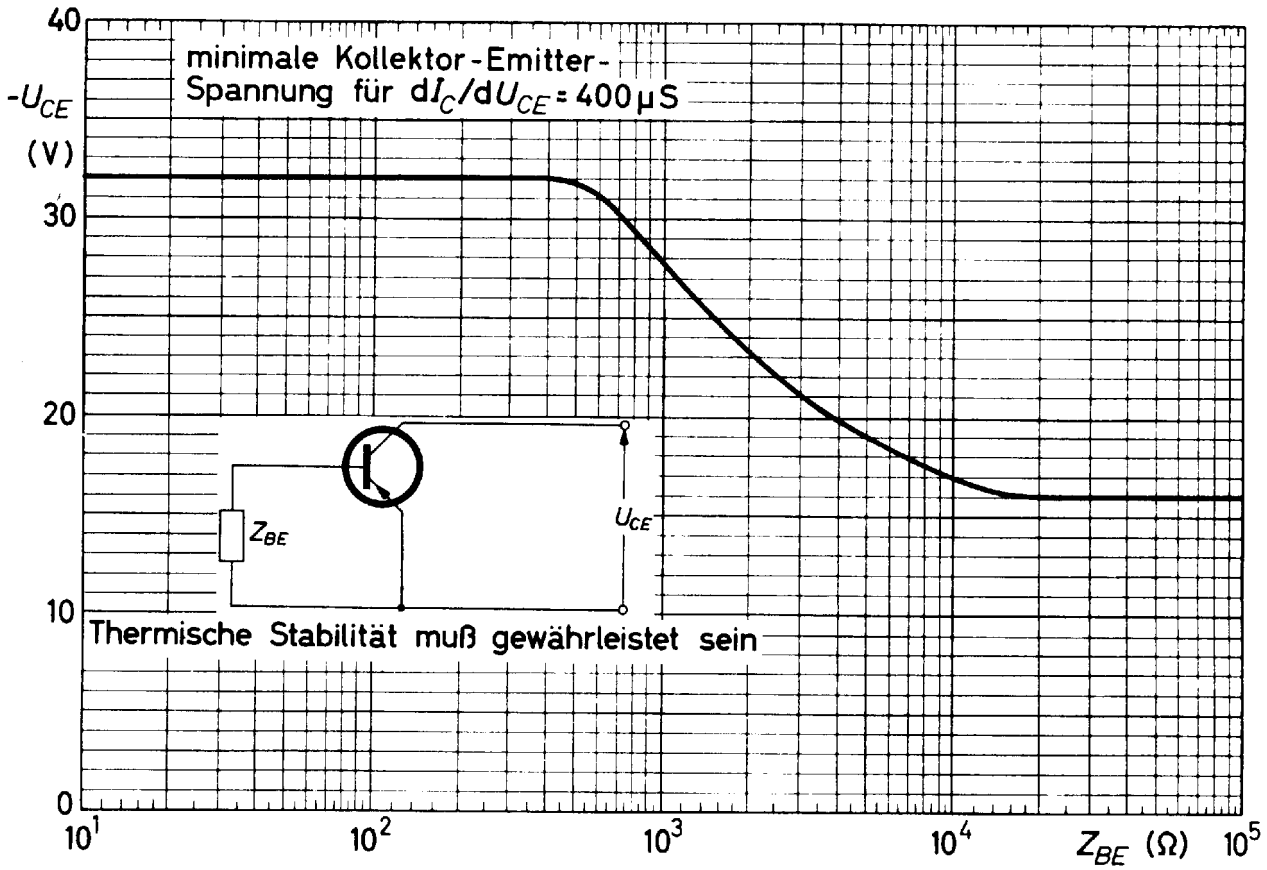

U_{CB} = 0
ϑ_j = 25°C
I_E (mA)
10
8
6
4
2
0
0
50
100
150
200 $-I_B$ (µA) 250

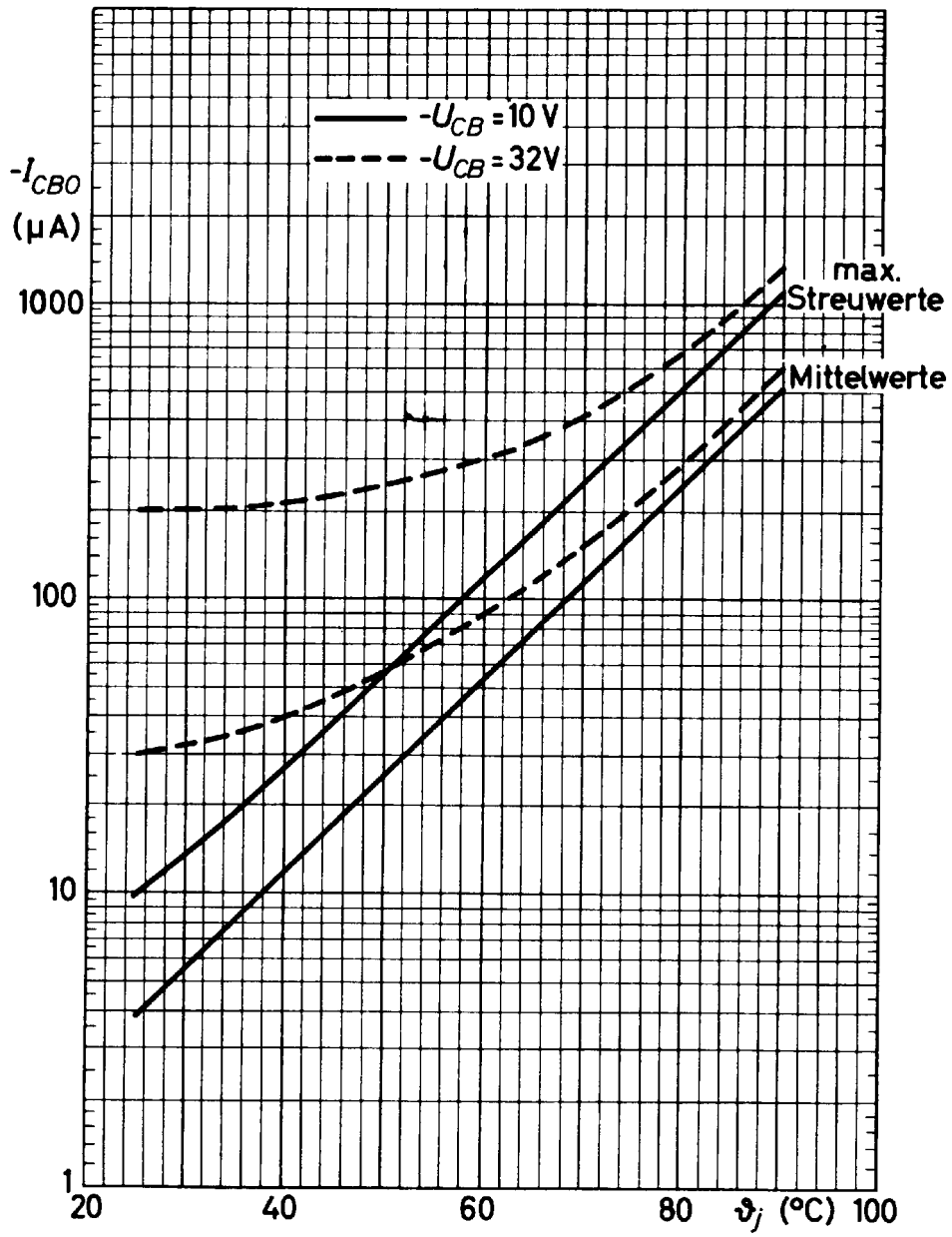

$-U_{CB}$ = 0
ϑ_j = 25°C
Mittelwert
Streuwerte
I_E (mA)
1400
1200
1000
800
600
400
200
0
0
5
10
15
20 $-I_B$ (mA) 25


I_E (mA)
10
8
6
4
2
0
$U_{CB} = 0$
$\vartheta_j = 25\,°C$
$\frac{\Delta(-U_{BE})}{\Delta\vartheta} = -2{,}3\,mV/grd$
$\frac{\Delta(-U_{BE})}{\Delta(-U_{CE})} = 0{,}4\,mV/V$
0
50
100
150 $-U_{BE}$ (mV) 200


I_E (mA)
1200
1000
800
600
400
200
0
$U_{CB} = 0$
$\vartheta_j = 25\,°C$
Mittelwert
Streuwerte
0
200
400
600 $-U_{BE}$ (mV) 800


β,B
150
100
50
0
β bei U_O =10V, R_L =16Ω
B bei $-U_{CE}$ =1V
0
100
200
300
400
$-I_C$ (mA) 500

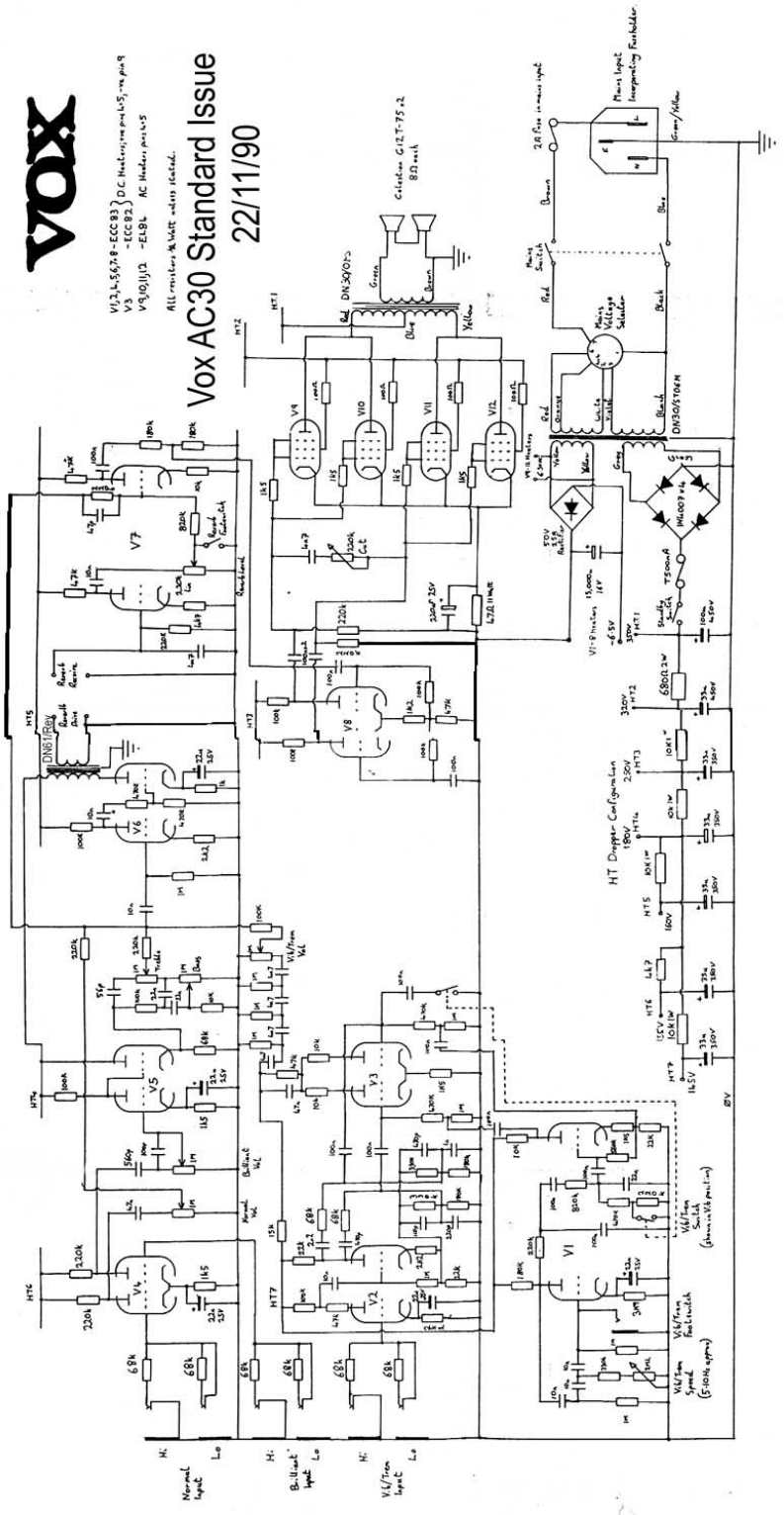

40
$-U_{CE}$
(V)
30
20
10
0
minimale Kollektor-Emitter-Spannung für dI_C/dU_{CE} = 400µS
Z_{BE}
U_{CE}
Thermische Stabilität muß gewährleistet sein
10^1
10^2
10^3
10^4
Z_{BE} (Ω)
10^5


$-U_{CB}$ = 10 V
$-U_{CB}$ = 32V
$-I_{CBO}$
(µA)
1000
100
10
1
20
40
60
80
ϑ_j (°C)
100
max.
Streuwerte
Mittelwerte

Appendix C

Vox AC30 Datasheet

Available, https://www.korguk.com/voxcircuits/, accessed 29/10/2018.


VOX
Vox AC30 Standard Issue
22/11/90

Appendix D

E-Appendices

Further supporting material is located at `http://bholmesqub.github.io/thesis/`.

- A video demonstrating the change in the behaviour of the diode clipper relative to input amplitude, visualising nonlinear impulse responses.
- MATLAB code for the root-finding algorithms in Chapter 3.
- Sound examples of the circuit models with different BJT component models in Chapter 4.
- MATLAB code for the multi-sine excitation signal described in Section 5.2.1.
- Sound examples of the Dallas Rangemaster Treble booster that is calibrated in Chapter 6.
- MATLAB code for the symbolic detection of redundancy of an RC circuit transfer function presented in Chapter 5.
- Sound examples of the identified tone stack and common-emitter amplifier in Chapter 7.

Bibliography

[1] S. Wilk, "What is the difference between analog and digital and why do I care for guitar effects pedals?." http://screaminfx.com/tech/analog-verse-digital-guitar-pedals.htm, 2014.

[2] V. Valimaki, S. Bilbao, J. O. Smith III, J. S. Abel, J. Pakarinen, and D. Berners, "Virtual analog effects," in *DAFX: Digital Audio Effects* (U. Zölzer, ed.), Chichester, U.K.: John Wiley & Sons, 2nd ed., 2011.

[3] L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, vol. 34, no. 1, pp. 1–12, 2010.

[4] L. W. Nagel and D. O. Pederson, "Simulation Program with Integrated Circuit Emphasis (SPICE)," in *Proceedings of the 16th Midwest Symposium on Circuit Theory*, (Waterloo, Canada), Apr. 1973.

[5] D. T.-M. Yeh, *Digital Implementation of Musical Distortion Circuits by Analysis and Simulation*. PhD thesis, Stanford University, 2009.

[6] O. Bogason and K. J. Werner, "Modeling Circuits with Operational Transconductance Amplifiers Using Wave Digital Filters," p. 8, 2017.

[7] W. R. Dunkel, M. Rest, K. J. Werner, M. J. Olsen, and J. O. S. Iii, "The Fender Bassman 5F6-A Family of Preamplifier Circuits—A Wave Digital Filter Case Study," p. 8, 2016.

[8] J. A. Seitchik, C. F. Machala, and P. Yang, "The determination of SPICE Gummel-Poon parameters by a merged optimization-extraction technique," in *Proc. of the 1989 Bipolar Circuits and Technology Meeting*, pp. 275–278, IEEE, 1989.

[9] T. Thiede, "PEAQ - The ITU Standardfor ObjectiveMeasuremenot f Perceived Audio Quality," *J. Audio Eng. Soc.*, vol. 48, no. 1, p. 27, 2000.

[10] F. Eichas and U. Zoelzer, "Guitar amplifier modeling—Perceptual evaluation of audio similarity," Mar. 2018.

[11] M. Holters and U. Zölzer, "Physical Modelling of a Wah-Wah Pedal as a Case Study for Application of the Nodal DK Method to Circuits with Variable Parts," in *Proc. of the 14th Internation Conference on Digital Audio Effects*, (Paris, France), pp. 31–35, Sept. 2011.

[12] A. Novak, L. Simon, P. Lotton, and J. Gilbert, "Chebyshev model and synchronized swept sine method in nonlinear audio effect modeling," in *Proc. 13th Int. Conference on Digital Audio Effects*, 2010.

[13] F. Eichas, S. Möller, and U. Zölzer, "Block-oriented modeling of distortion audio effects using iterative minimization," in *Proceedings of the 18th International Conference on Digital Audio Effects*, (Trondheim, Norway), Dec. 2015.

[14] D. T. Yeh, "Automated Physical Modeling of Nonlinear Audio Circuits for Real-Time Audio Effects—Part II: BJT and Vacuum Tube Examples," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 1207–1216, May 2012.

[15] K. Dempwolf and U. Zölzer, "Discrete State-Space Model of the Fuzz-Face," in *Proceedings of Forum Acusticum*, (Aalborg, Denmark), European Acoustics Association, June 2011.

[16] F. Eichas, M. Fink, M. Holters, and U. Zölzer, "Physical Modeling of the MXR Phase 90 Guitar Effect Pedal," in *Proc. of the 17 Th Int. Conference on Digital Audio Effects*, (Erlangen, Germany), pp. 153–156, Sept. 2014.

[17] J. Macak, *Real-Time Digital Simulation of Guitar Amplifiers as Audio Effects*. PhD thesis, Brno University of Technology, 2012.

[18] D. T. Yeh, J. Abel, and J. O. Smith, "Simulation of the diode limiter in guitar distortion circuits by numerical solution of ordinary differential equations," *Proceedings of the Digital Audio Effects (DAFx'07)*, pp. 197–204, 2007.

[19] R. C. D. Paiva, S. D'Angelo, J. Pakarinen, and V. Valimaki, "Emulation of Operational Amplifiers and Diodes in Audio Distortion Circuits," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, pp. 688–692, Oct. 2012.

[20] A. Falaize and T. Hélie, "Passive Guaranteed Simulation of Analog Audio Circuits: A Port-Hamiltonian Approach," *Applied Sciences*, vol. 6, p. 273, Sept. 2016.

[21] J. Macak and J. Schimmel, "Simulation of a vacuum-tube push-pull guitar power amplifier," in *Proc. of the 14th Internation Conference on Digital Audio Effects*, (Paris, France), p. 4, 2011.

[22] M. Holters and U. Zölzer, "Circuit Simulation with Inductors and Transformers Based on the Jiles-Atherton Model of Magnetization," p. 6, 2016.

[23] Z. Jingjie and J. O. Smith III, "Real-Time Wave Digital Simulation of Casceded Vacuum Tube Amplifiers using Modified Block-Wise Method," in *Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18), Aveiro, Portugal, September 4–8, 2018*, (Aveiro, Portugal), p. 8, Sept. 2018.

[24] J. Macak and J. Schimmel, "Real-time guitar tube amplifier simulation using an approximation of differential equations," in *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx'10)*, 2010.

[25] N. Koren, "Improved vacuum tube models for SPICE, Part 1." http://www.normankoren.com/Audio/Tubemodspice_article.html, 2003. [Online] - Available: http://www.normankoren.com/Audio/Tubemodspice_article.html accessed 14/11/2014.

[26] G. C. Cardarilli, "Improved large-signal model for vacuum triodes," in *IEEE Int. Symp. Circuits Syst.*, (Taipei, Taiwan), May 2009.

[27] K. Dempwolf, M. Holters, and U. Zölzer, "A triode model for guitar amplifier simulation with individual parameter fitting," in *Audio Engineering Society Convention 131*, Audio Engineering Society, 2011.

[28] O. Kroening, "Planning and Construction of an Automated Measuring Object for Electron Tubes," Feb. 2012.

[29] K. Dempwolf and U. Zölzer, "A physically-motivated triode model for circuit simulations," in *14th International Conference on Digital Audio Effects DAFx*, vol. 11, 2011.

[30] D. T. Yeh, J. S. Abel, and J. O. Smith, "Automated Physical Modeling of Nonlinear Audio Circuits For Real-Time Audio Effects; Part I: Theoretical Development," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 728–737, May 2010.

[31] K. J. Werner, V. Nangia, J. O. Smith, and J. S. Abel, "A general and explicit formulation for wave digital filters with multiple/multiport nonlinearities and complicated topologies," in *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, (New Paltz, NY, USA), pp. 1–5, IEEE, Oct. 2015.

[32] C.-W. Ho, A. E. Ruehli, and P. A. Brennan, "The modified nodal approach to network analysis," *Circuits and Systems, IEEE Transactions on*, vol. 22, no. 6, pp. 504–509, 1975.

[33] G. Borin, G. De Poli, and D. Rocchesso, "Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems," *Speech and Audio Processing, IEEE Transactions on*, vol. 8, no. 5, pp. 597–605, 2000.

[34] D. Cole, B. Morgan, and D. Titterington, "Determining the parametric structure of models," *Mathematical Biosciences*, vol. 228, pp. 16–30, Nov. 2010.

[35] C. Gnegy and K. J. Werner, "Digitizing the Ibanez Weeping Demon Wah Pedal," p. 8, 2015.

[36] K. J. Werner, J. S. Abel, and J. O. Smith, "More Cowbell: A Physically-Informed, Circuit-Bendable, Digital Model of the TR-808 Cowbell," *Los Angeles*, p. 10, 2014.

[37] D. T. Yeh and J. O. Smith, "Discretization of the '59 Fender Bassman tone stack," in *Proc. of the Int. Conf. on Digital Audio Effects (DAFx-06)*, pp. 18–20, 2006.

[38] A. Fettweis, "Digital filters related to classical structures," *AEU: Archive für Elektronik und Übertragungstechnik*, vol. 25, pp. 78–89, Feb. 1971.

[39] A. Fettweis, "Wave Digital Filters: Theory and Practice," *Proceeding of the IEEE*, vol. 74, Feb. 1986.

[40] G. de Sanctis and A. Sarti, "Virtual Analog Modeling in the Wave-Digital Domain," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 715–727, May 2010.

[41] M. Karjalainen and J. Pakarinen, "Wave Digital Simulation of a Vacuum-Tube Amplifier.," in *ICASSP (5)*, pp. 153–156, 2006.

[42] J. Pakarinen and M. Karjalainen, "Enhanced Wave Digital Triode Model for Real-Time Tube Amplifier Emulation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 738–746, May 2010.

[43] T. Schwerdtfeger and A. Kummert, "A multidimensional signal processing approach to Wave Digital Filters with topology-related delay-free loops," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 389–393, May 2014.

[44] T. Schwerdtfeger and A. Kummert, "A multidimensional approach to wave digital filters with multiple nonlinearities," in *2014 22nd European Signal Processing Conference (EUSIPCO)*, p. 5, 2014.

[45] T. Schwerdtfeger and A. Kummert, "Newton's method for modularity-preserving multidimensional wave digital filters," in *2015 IEEE 9th International Workshop on Multidimensional (nD) Systems (nDS)*, pp. 1–6, Sept. 2015.

[46] D. Franken, J. Ochs, and K. Ochs, "Generation of wave digital structures for connection networks containing ideal transformers," in *Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03.*, vol. 3, (Bangkok, Thailand), pp. III–240–III–243, IEEE, 2003.

[47] D. Franken, J. Ochs, and K. Ochs, "Generation of wave digital structures for networks containing multiport elements," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, pp. 586–596, Mar. 2005.

[48] K. J. Werner, V. Nangia, J. O. S. Iii, and J. S. Abel, "Resolving Wave Digital Filters with Multiple/Multiport Nonlinearities," p. 8, 2015.

[49] K. J. Werner, J. O. Smith III, and J. S. Abel, "Wave digital filter adaptors for arbitrary topologies and multiport linear elements," in *Proceedings of the 18th International Conference on Digital Audio Effects*, (Trondheim, Norway), Dec. 2015.

[50] M. J. Olsen, K. J. Werner, and J. O. S. IIIcolorlinks, "Resolving Grouped Nonlinearities in Wave Digital Filters Using Iterative Techniques," p. 8, 2016.

[51] K. J. Werner, M. J. Olsen, M. Rest, and J. D. Parker, "Generalizing Root Variable Choice in Wave Digital Filters with Grouped Nonlinearities," p. 8, 2017.

[52] K. J. Werner, A. Bernardini, J. O. S. Iii, and A. Sarti, "Modeling Circuits with Arbitrary Topologies and Active Linear Multiports using Wave Digital Filters," *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS*, p. 14, June 2018.

[53] A. Falaize-Skrzek and T. Hélie, "Simulation of an analog circuit of a wah pedal: A port-Hamiltonian approach," in *Audio Engineering Society Convention 135*, Audio Engineering Society, 2013.

[54] A. Falaize and T. Helie, "Passive Simulation of Electrodynamic Loudspeakers for Guitar Amplifiers: A Port-Hamiltonian Approach," in *Proc. of the International Conference On Noise and Vibration Engineering (ISMA)*, (Le Mans, France), 2014.

[55] A. Falaize and T. Hélie, "Passive simulation of the nonlinear port-Hamiltonian modeling of a Rhodes Piano," *Journal of Sound and Vibration*, vol. 390, pp. 289–309, Mar. 2017.

[56] R. Muller and T. Hélie, "Trajectory Anti-aliasing on Guaranteed-passive Simulation of Nonlinear Physical Systems," p. 8, 2017.

[57] R. Müller and T. Hélie, "Power-balanced Modelling of Circuits as Skew Gradient Systems," in *Proceedings of the 21 St International Conference on Digital Audio Effects (DAFx-18), Aveiro, Portugal, September 4–8, 2018*, p. 8, 2018.

[58] J. Macak and J. Schimmel, "Real-Time Guitar Preamp Simulation Using Modified Blockwise Method and Approximations," *EURASIP Journal on Advances in Signal Processing*, 2011.

[59] J. Macak, "Guitar Preamp Simulation Using Connection Currents," in *Proceedings of the International Conference on Digital Audio Effects (DAFx-16), Maynooth, Ireland*, 2013.

[60] M. Holters and U. Zölzer, "A generalized method for the derivation of nonlinear state-space models from circuit schematics," in *23rd European Signal Processing Conference (EUSIPCO), 2015*, 2015.

[61] A. Novak, P. Lotton, and L. Simon, "Synchronized Swept-Sine: Theory, Application, and Implementation," *Journal of the Audio Engineering Society*, vol. 63, pp. 786–798, Nov. 2015.

[62] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio Engineering Society Convention 108*, Audio Engineering Society, 2000.

[63] A. Novak, L. Simon, P. Lotton, and F. Kadlec, "Modeling of nonlinear audio systems using swept-sine signals: Application to audio effects," in *Proc. of the 12th Int. Conference on Digital Audio Effects (DAFx-09)*, pp. 1–4, 2009.

[64] L. Tronchin and V. L. Coli, "Further Investigations in the Emulation of Nonlinear Systems with Volterra Series," *Journal of the Audio Engineering Society*, vol. 63, pp. 671–683, Oct. 2015.

[65] T. Schmitz and J.-J. Embrechts, "Hammerstein Kernels Identification by Means of a Sine Sweep Technique Applied to Nonlinear Audio Devices Emulation," *Journal of the Audio Engineering Society*, vol. 65, pp. 696–710, Sept. 2017.

[66] T. Schmitz and J.-J. Embrechts, "Real Time Emulation of Parametric Guitar Tube Amplifier with Long Short Term Memory Neural Network," in *Computer Science & Information Technology*, pp. 149–157, Academy & Industry Research Collaboration Center (AIRCC), Apr. 2018.

[67] C. Kemper, "Musical instrument with acoustic transducer," Aug. 2014. US Patent 8,796,530.

[68] F. Eichas, S. Möller, and U. Zölzer, "Block-oriented Gray Box Modeling of Guitar Amplifiers," p. 8, 2017.

[69] W. Warren, "Builder Profile: Klon's Bill Finnegan." https://www.premierguitar.com/articles/20210-builder-profile-klons-bill-finnegan, Jan. 2014.

[70] J. Macak, J. Schimmel, and M. Holters, "Simulation of fender type guitar preamp using approximation and state-space model," in *Proceedings of the 12th International Conference on Digital Audio Effects, York, UK*, 2012.

[71] C. T. Kelley, *Solving Nonlinear Equations with Newton's Method*. Fundamentals of Algorithms, Philadelphia: Society for Industrial and Applied Mathematics, 2003.

[72] V. Chatziioannou, S. Schmutzhard, and S. Bilbao, "On Iterative Solutions for Numerical Collision Models," p. 8, 2017.

[73] S. D'Angelo, J. Pakarinen, and V. Valimaki, "New Family of Wave-Digital Triode Models," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 2, pp. 313–321, 2013.

[74] C. G. Broyden, "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Mathematics of Computation*, vol. 19, p. 577, Oct. 1965.

[75] G. Alefeld, "On the Convergence of Halley's Method," *The American Mathematical Monthly*, vol. 88, p. 530, Aug. 1981.

[76] A. A. M. Cuyt and L. B. Rall, "Computational Implementation of the Multivariate Halley Method for Solving Nonlinear Systems of Equations," *ACM Trans. Math. Softw.*, vol. 11, pp. 20–36, Mar. 1985.

[77] R. L. Burden and J. D. Faires, *Numerical Analysis*. PWS Publishers, 3rd ed. ed., 1985.

[78] T. Banwell and A. Jayakumar, "Exact analytical solution for current flow through diode with series resistance," *Electronics Letters*, vol. 36, pp. 291–292, Feb. 2000.

[79] A. Ushida and Y. Inoue, "An Efficient Algorithm for Finding Multiple DC Solutions Based on the SPICE-Oriented Newton Homotopy Method," vol. 21, no. 3, p. 12, 2002.

[80] A. Vladimirescu, *The SPICE Book*. New York: J. Wiley, 1994.

[81] T. Minka, "The Lightspeed Matlab Toolbox." http://research.microsoft.com/en-us/um/people/minka/software/lightspeed/. [Online]. Available: http://research.microsoft.com/en-us/um/people/minka/software/lightspeed/ - accessed 22/04/2015.

[82] J. L. Hennessy and D. A. Patterson, "Instruction-Level Parallelism: Concepts and Challenges," in *Computer Architecture: A Quantitative Approach*, Waltham, MA: Morgan Kaufmann/Elsevier, 5th ed., 2012.

[83] B. Holmes and M. van Walstijn, "Improving the robustness of the iterative solver in state-space modelling of guitar distortion circuitry," in *Proc. of the 18th International Conference on Digital Audio Effects*, (Trondheim, Norway), pp. 49–56, Dec. 2015.

[84] J. J. Ebers and J. L. Moll, "Large-signal behavior of junction transistors," *Proceedings of the IRE*, vol. 42, no. 12, pp. 1761–1772, 1954.

[85] H. K. Gummel and H. C. Poon, "An integral charge control model of bipolar transistors," *Bell System Technical Journal*, vol. 49, no. 5, pp. 827–852, 1970.

[86] C. C. McAndrew, J. A. Seitchik, D. F. Bowers, M. Dunn, M. Foisy, I. Getreu, M. McSwain, S. Moinian, J. Parker, D. J. Roulston, *et al.*, "VBIC95, the vertical bipolar inter-company model," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 10, pp. 1476–1483, 1996.

[87] R. Van der Toorn, J. C. J. Paasschens, and W. J. Kloosterman, "The Mextram bipolar transistor model," *Delft University of Technology, Technical report*, 2008.

[88] M. Dregni, "The Dallas Rangemaster," *Vintage Guitar*, Sept. 2014.

[89] R. C. D. Paiva and V. V. L. Ki, "Acoustics and Modeling of Pickups," *J. Audio Eng. Soc.*, vol. 60, no. 10, p. 15, 2012.

[90] M. Dregni, "The Arbiter Fuzz Face," *Vintage Guitar*, Aug. 2012.

[91] R. C. Jaeger, *Microelectronic Circuit Design*. New York, NY: McGraw-Hill, a business unit of The McGraw-Hill Companies, Inc, fifth edition ed., 2015.

[92] I. Getreu, *Modeling the Bipolar Transistor*. Tektronix, 1976.

[93] J. M. Miller, "Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit," *Scientific Papers of the Bureau of Standards*, vol. 15, no. 351, pp. 367–385, 1920.

[94] F. Sischka, "Gummel-Poon Bipolar Model: Model description, parameter extraction.," *Agilent Technologies*, 2001.

[95] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder–Mead simplex method in low dimensions," *SIAM Journal on optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[96] R. H. Byrd, J. C. Gilbert, and J. Nocedal, "A Trust Region Method Based on Interior Point Techniques for Nonlinear Programming," *Mathematical Programming*, vol. 89, no. 1, pp. 149–185, 2000.

[97] B. Holmes, M. Holters, and M. van Walstijn, "Comparison of Germanium Bipolar Junction Transistor Models for Real-time Circuit Simulation," p. 8, 2017.

[98] K. Wyatt, "Resistors aren't resistors," Oct. 2013.

[99] M. Schroeder, "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (Corresp.)," *Information Theory, IEEE transactions on*, vol. 16, no. 1, pp. 85–89, 1970.

[100] A. Saltelli, ed., *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. Hoboken, NJ: Wiley, 2004.

[101] M. D. Morris, "Factorial Sampling Plans for Preliminary Computational Experiments," *Technometrics*, vol. 33, pp. 161–174, May 1991.

[102] H. Sohier, H. Piet-Lahanier, and J.-L. Farges, "Analysis and optimization of an air-launch-to-orbit separation," *Acta Astronautica*, vol. 108, pp. 18–29, Mar. 2015.

[103] F. Campolongo, J. Cariboni, and A. Saltelli, "An effective screening design for sensitivity analysis of large models," *Environmental Modelling & Software*, vol. 22, pp. 1509–1518, Oct. 2007.

[104] F. Pianosi, F. Sarrazin, and T. Wagener, "A Matlab toolbox for Global Sensitivity Analysis," *Environmental Modelling & Software*, vol. 70, pp. 80–85, Aug. 2015.

[105] M. D. McKay, R. J. Beckman, and W. J. Conover, "Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, vol. 21, pp. 239–245, May 1979.

[106] IEC, "IEC 60063:2015 - Preferred number series for resistors and capacitors," tech. rep., Mar. 2015.

[107] O. Semiconductor, "2N3906 - General Purpose Transistors, PNP Silicon," p. 7, 2010.

[108] Vishay, "1n4148 - Small signal fast switching diodes," tech. rep., 2017.

[109] O. Semiconductor, "1N4001 - Axial-Lead Glass Passivated Standard Recovery Rectifiers," p. 7, June 2018.

[110] B. Holmes and M. van Walstijn, "Physical model parameter optimisation for calibrated emulation of the Dallas Rangemaster Treble Booster guitar pedal," in *Proc. of the 19th Internation Conference on Digital Audio Effects*, (Brno, Czech Republic), pp. 47–54, Sept. 2016.

[111] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional, 1989.

[112] J. Bandler, R. Biernacki, Shao Hua Chen, P. Grobelny, and R. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 12, pp. 2536–2544, Dec./1994.